

(19)



Europäisches Patentamt  
European Patent Office  
Office européen des brevets



(11)

EP 0 753 584 A1

(12)

## EUROPEAN PATENT APPLICATION

(43) Date of publication:  
15.01.1997 Bulletin 1997/03

(51) Int. Cl.<sup>6</sup>: C12Q 1/68, C07D 209/62,  
C09B 23/00

(21) Application number: 96250149.0

(22) Date of filing: 05.07.1996

(84) Designated Contracting States:  
DE FR GB SE

(30) Priority: 11.07.1995 US 500691

(71) Applicant: LI-COR, INC.  
Lincoln, Nebraska 68504 (US)

(72) Inventors:  
• Patoay, Gabor  
Conyers, Georgia 30208 (US)

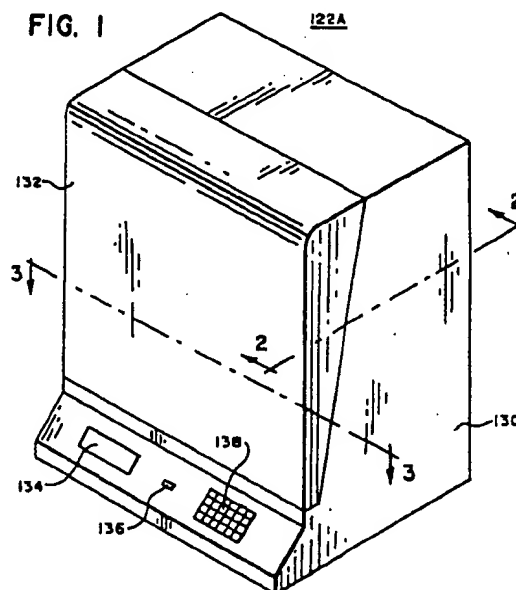
• Narayanan, Narasimhachari  
Lincoln, Nebraska 68516 (US)  
• Brumbaugh, John Albert  
Estes Park, Colorado 80517 (US)  
• Middendorf, Lyle Richard  
Lincoln, Nebraska 68505 (US)

(74) Representative: UEXKÜLL & STOLBERG  
Patentanwälte  
Beselerstrasse 4  
22607 Hamburg (DE)

(54) Sequencing near infrared and infrared fluorescence labeled DNA for detecting using laser diodes and suitable labels therefor

(57) To sequence DNA automatically, DNA marked with far infrared, near infrared, or infrared fluorescent dyes are electrophoresed in a plurality of channels through a gel electrophoresis slab or capillary tubes wherein the DNA samples are resolved in accordance with the size of DNA fragments in the gel electrophoresis slab or capillary tubes into fluorescently marked DNA bands. The separated samples are scanned photoelectrically with a laser diode and a sensor, wherein the laser scans with scanning light at a wavelength within the absorbance spectrum of said fluorescently marked DNA samples and light is sensed at the emission wavelength of the marked DNA.

FIG. 1



EP 0 753 584 A1

## Description

This invention relates to the sequencing of fluorescence labeled DNA and the detecting of the DNA after irradiation by light from a laser, and suitable labels therefor.

In one class of techniques for sequencing DNA, identical strands of DNA are marked with a fluorescent dye. The strands are marked by attaching specially synthesized fluorescent oligonucleotide primers or probes to the strands of DNA, or by attaching the fluorescent dye directly to the strands.

The strands are separated into four aliquots. The strands in a given aliquot are either individually cleaved at or synthesized to any base belonging to only one of the four base types, which are adenine, guanine, cytosine, and thymine (hereinafter A, G, C and T). The adenine, guanine, cytosine, and thymine terminated strands are then electrophoresed for separation and the separated strands are irradiated by a laser and the emission from the fluorescent dye detected. The rate of electrophoresis indicates the DNA sequence.

In a prior art sequencing technique of this class, the fluorescent dyes used as markers have their maximum emission spectra in the visible range, the DNA is subject to irradiation in the visible spectra, and visible spectra detectors and light sources are used. Generally photomultiplier tubes (PMT) are used for detection.

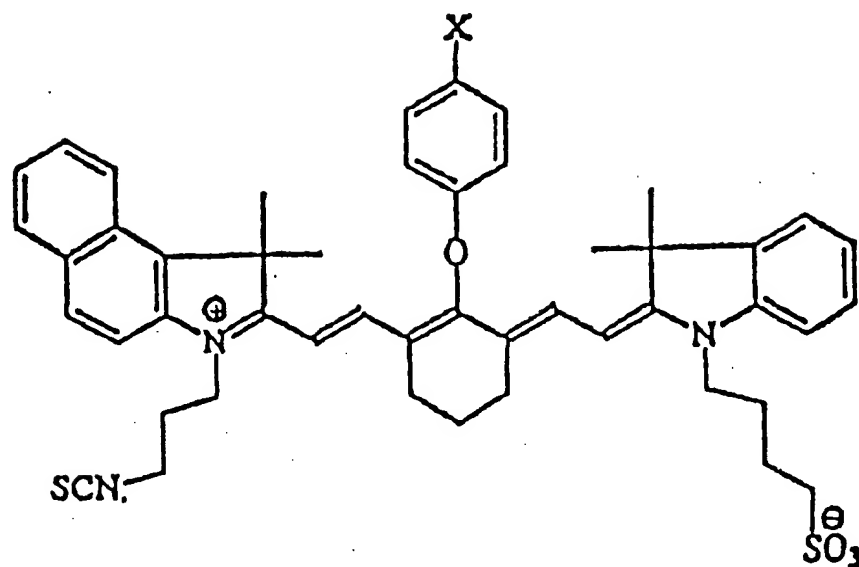
The prior art techniques for DNA sequencing have several disadvantages such as: (1) because the dyes have their emission spectra in the visible region of light spectrum, the lasers used to excite the fluorescent markers, and under some circumstances, the detectors for the light tend to be expensive; and (2) they are relatively noisy due to the background interferences by biomolecules.

Cyanine dyes are known to absorb far-red (600-700 nm) and near-infrared (700-1200 nm) light and techniques for the synthesis of derivatives of the cyanine dyes are known. However, obtaining chromophores with absorbance and emission bands that reduce the effect of background noise during gel electrophoresis apparatus when irradiating with a diode laser scanning has been difficult. Other dyes, not described herein, exhibit absorbance in wavelengths greater than 1200 nm (infrared) and would also provide discrimination against background noise.

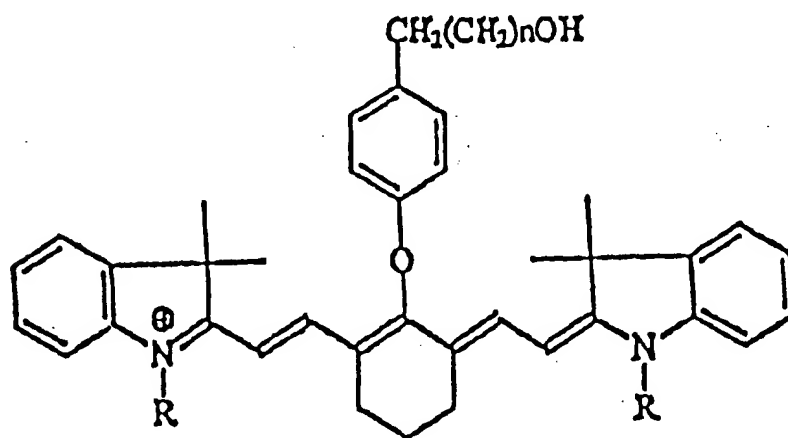
Suitable types of cyanine dyes include heptamethine cyanine dyes. Cyanine dyes have traditionally been synthesized by a condensation reaction between a heterocyclic base containing an activated methyl group and an unsaturated bisaldehyde or its equivalent, usually as Schiff base in the presence of a catalyst. Sodium acetate has most frequently been used as a catalyst. In addition to ethanol, solvents such as acetic acid and/or acetic anhydride have also been commonly used as in the synthesis of heptamethine pyrylium dyes. This procedure suffers from several disadvantages, such as for example: (1) the purification of the product is very difficult because of the side products due to aniline; (2) the use of a catalyst interferes with the purity of the product and warrants repeated purification; (3) the reaction is generally faster and cannot be employed for the synthesis of nonsymmetric dyes in one pot; and (4) the scaling up of the reaction products to larger gram quantities leads to several additional problems resulting in poor quality and yield.

It is a task of the invention to provide novel dyes that fluoresce in the far red, near infrared, or infrared region in selected wavelengths when attached to oligonucleotides and that have sufficient quantum yield to make detection feasible for use especially in probes or primers.

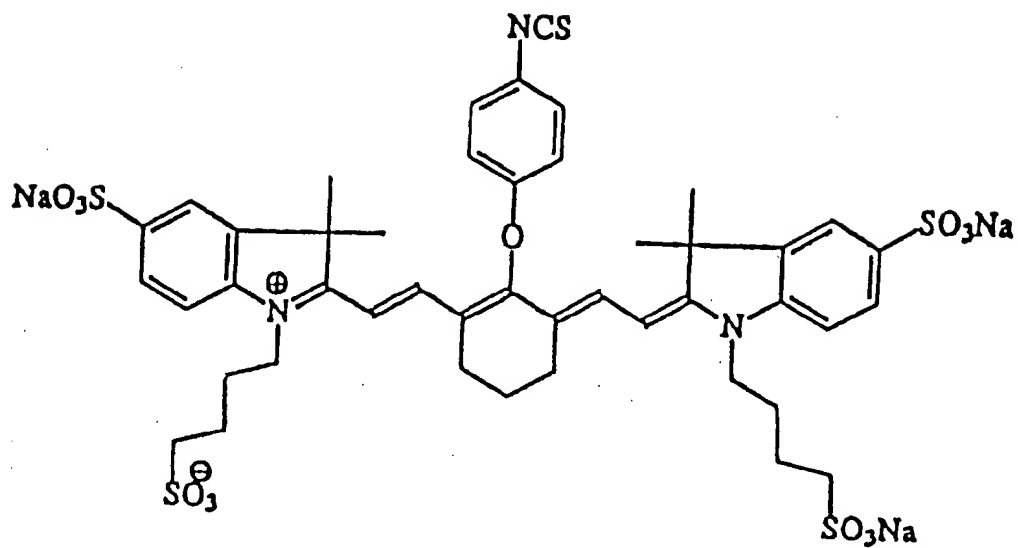
To accomplish this above task, a method of identifying strands of DNA comprises marking the strands with fluorescent labels that emit light, irradiating the strands with light, and detecting the light emitted from the fluorescent labels, characterized in that the fluorescent labels emit light in a region of wavelengths including at least one wavelength within one of a far red, near infrared, and infrared regions wherein the fluorescent label includes a chromophore having the formula:



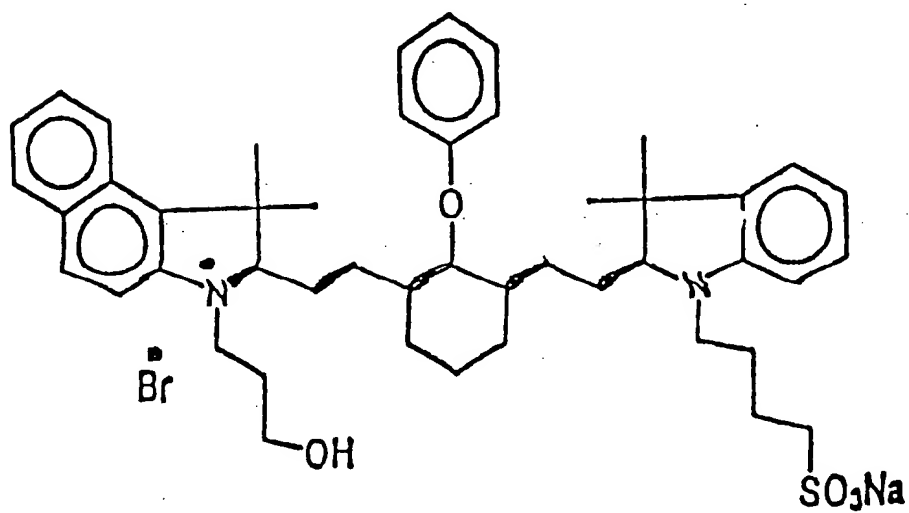
where X is H or OCH<sub>3</sub>;  
or the formula:



where R is ethyl or 4-sulfonatobutyl and n is either 1 or 2.  
or the formula



or the formula



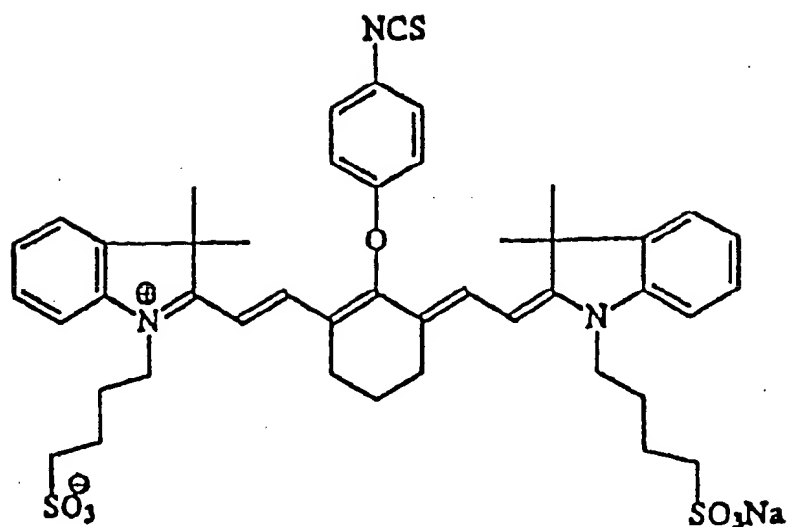
or the formula

5

10

15

20



and the irradiating light has a wavelength within one of the far red, near infrared, and infrared regions.

25

Advantageously, the irradiating is with one of the far red, near infrared, and infrared laser diode light sources. In this method, DNA samples marked with fluorescent dye having absorbance and fluorescence maxima at one of the far red, near infrared, and infrared wavelengths when combined with the DNA are applied at a plurality of locations in a gel electrophoresis slab or set of capillaries for electrophoresing in a plurality of channels through the gel electrophoresis slab or set of capillaries establishing electrical potential across said gel electrophoresis slab or set of capillaries wherein DNA samples are resolved in accordance with the size of DNA fragments in said gel electrophoresis slab or set of capillaries into fluorescently marked DNA bands; and irradiating the separated samples photo-electrically while they are in the slab or set of capillaries with a laser diode and a sensor wherein the laser scans with far red, near infrared, or infrared scanning light at the wavelength within an absorbance spectrum of said fluorescently marked DNA samples and sensing light at the emission wavelength of the marked DNA.

35

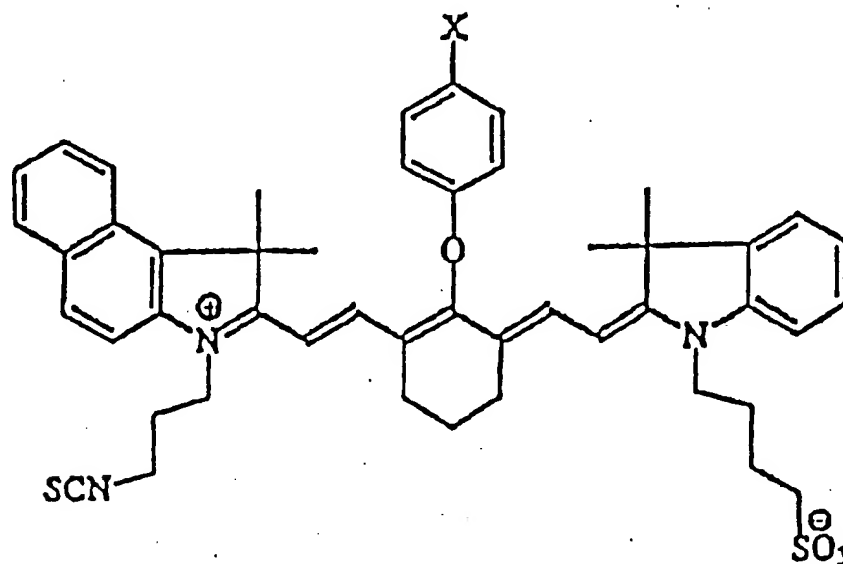
A dye characterized by one of the formula:

40

45

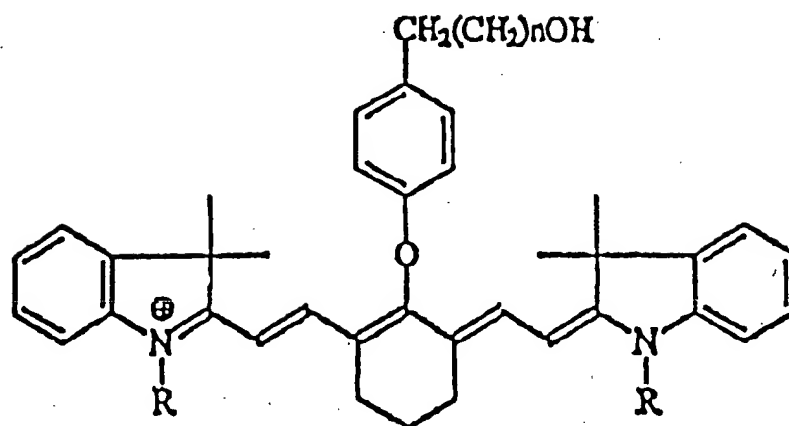
50

55

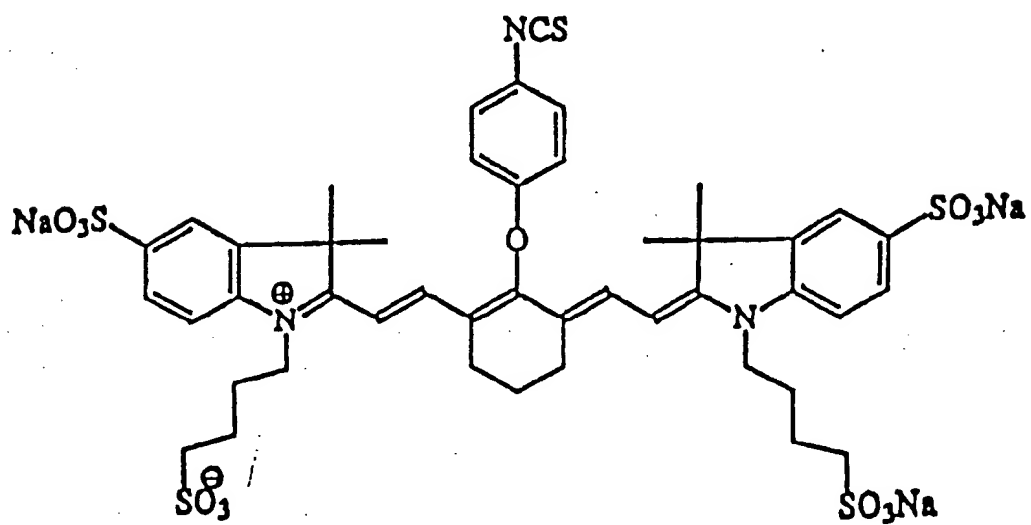


where X is H or OCH<sub>3</sub>

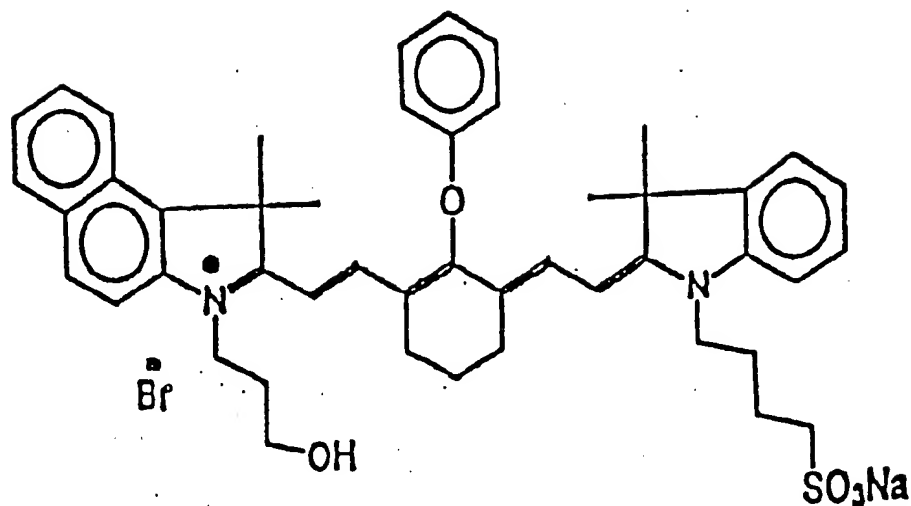
or



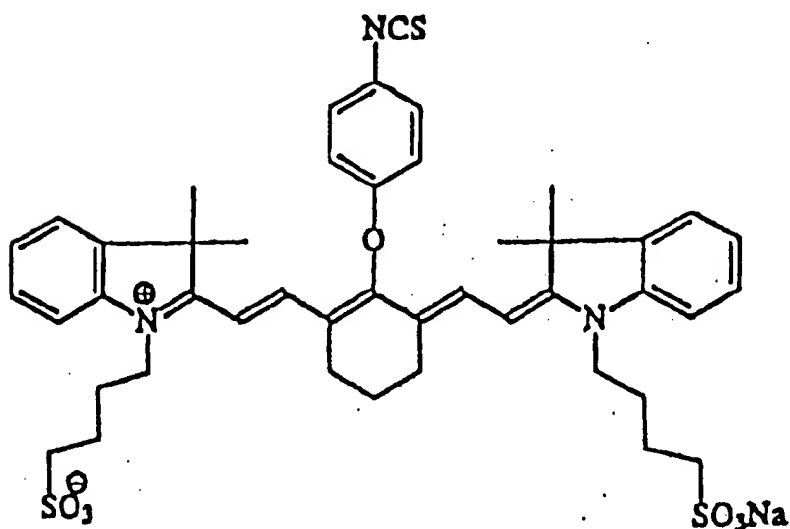
where R is ethyl or sulfonatobutyl and n is either 1 or 2;  
or



or



or



From the above summary, it can be understood that the technique for the sequencing of fluorescence labeled DNA of this invention have several advantages, such as: (1) because the dyes have their emission spectra in the far red, near infrared, or infrared light spectrum, small inexpensive diode lasers may be used; and (2) this wavelength region is characterized by relatively low background fluorescence in glass, and therefore, less noise in the received signal.

The above noted and other features of the invention will be better understood from the following detailed description when considered with reference to the accompanying drawings in which:

FIG. 1 is a perspective view of an embodiment of sequencer usable in the invention;

FIG. 2 is a sectional view taken through lines 2-2 of FIG. 1;

FIG. 3 is a sectional view of a portion of FIG. 1 taken through lines 3-3;

FIG. 4 is an exploded perspective view of a portion of the embodiment of FIG. 2;

FIG. 5 is an enlarged view, partly broken away, of a portion of the embodiment of FIG. 2;

FIG. 6 is a block diagram of a circuit that may be used for coordination of a sensor, scanner drive, and laser used; and

FIG. 7 is an embodiment of electrophoresis apparatus that may be used instead of the gel sandwich used in the

embodiments of FIGS. 1-6.

The sequencing of far red, near infrared, or infrared fluorescence labeled DNA and the detection of the DNA after irradiation by far red, near infrared, or infrared light from a laser diode is accomplished using a far red, near infrared, or infrared label prepared for this purpose and either directly attached to the DNA or attached to probes or primers that will be attached to the DNA. In this specification the word "infrared" will be used at times to include far red wavelengths (600-700 nm), near infrared (700-1200 nm) and infrared (1200-4000 nm). The strands of DNA are continuously electrophoresed and identified for any of several purposes, such as for example: (1) DNA sequencing; and (2) analysis of strands varying in length as prepared by such techniques as restriction enzyme cutting or polymerase chain reaction (PCR).

The strands are marked with fluorescent labels that have their maximum fluorescence and their maximum absorbance at wavelengths of light in the far red, near infrared, or infrared region. The strands are irradiated with light in the far red, near infrared, or infrared region from a laser diode and the light emitted from the fluorescent labels is detected and used to obtain information about the DNA strands. The detector includes a light sensor which is preferably an avalanche photodiode sensitive to the infrared light emission of the marker. It may include a filtering system having a pass band suitable for passing selectively the optimum emission of the fluorescent marker to the light sensor.

To mark the DNA strand, a dye is prepared having the desired properties or a known dye is modified. In the preferred embodiment a novel dye having the preferred spectrum, high absorption and fluorescence properties, and at least one reactive group enabling coupling to biomolecules such as DNA, proteins, and antibodies is synthesized.

The dye is synthesized, modified from a known dye or selected to have an absorbance band and an emission band within a region encompassing the far red, near infrared, or infrared region when attached to a probe, primer, oligonucleotide or other molecule. The dye should provide high quantum yield in an optical band selected to reduce background noise. The preferred dyes for many applications calling for the labelling of biomolecules are cyanine dyes having an NCS group on the dye that may react with the amino group of the biomolecule to form a thiourea linkage.

In the preferred embodiment, cyanine dyes are synthesized. The preferred dyes are heptamethine cyanines which efficiently absorb light having wavelengths in the region of 750 to 820 nm (nanometers) (maximum absorbance wavelength). This wavelength is suitable for reducing background fluorescence in DNA sequencing and corresponds to the radiation wavelength of the GaAlAs diode laser which is 780-800 nm. The GaAlAs diode is used for irradiating the gel electrophoresis sandwich, column, or capillary used for DNA sequencing. Formulas 1-13 are typical synthesized or proposed dyes and formula 14 is a suitable modified dye.

Formula 1 shows synthesized cyanine dyes having NCS (isothiocyanate) as a reactive group for attachment of a biomolecule. In this embodiment, when X is H, the maximum absorbance wavelength is 787 nm in methanol and 801 nm in DMSO, and the maximum emission wavelength is 807 nm in methanol and 818 nm in DMSO. When X is -OCH<sub>3</sub>, the maximum absorbance wavelength is 786 nm in methanol and the maximum emission wavelength is 806 nm in methanol. In both cases, the quantum yield is greater than 15%.

Formula 2 shows synthesized cyanine dyes which have a high quantum yield in methanol around 35%. When R is ethyl or sulfonatobutyl and n is either 1 or 2, the maximum absorbance wavelength is between 762 and 778 nm. Depending on the solvent, the maximum emission wavelength is between 782 and 800 nm.

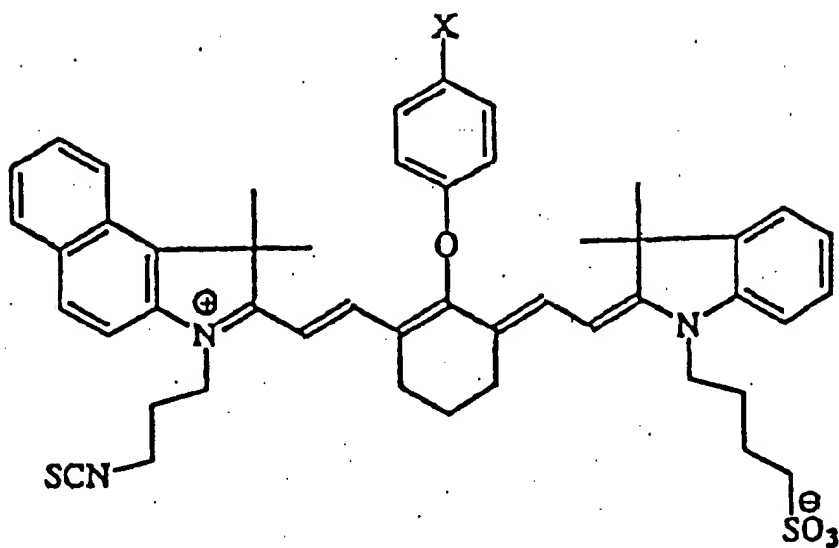
In the synthesized dye represented by formula 3, the maximum absorbance wavelength is between 773 and 778 nm, depending on the solvent and the maximum emission wavelength is between 789 and 806 nm, depending upon the solvent. The quantum yield is between 25% and 35%, depending upon the solvent.

In the molecule shown by formula 4, the dye contains a functional hydroxy group which either can be converted to NHS ester to permit coupling to DNA strands for the purposes of sequencing DNA or directly phosphitylated for use as a dye-labeled amidite for use in DNA synthesis.

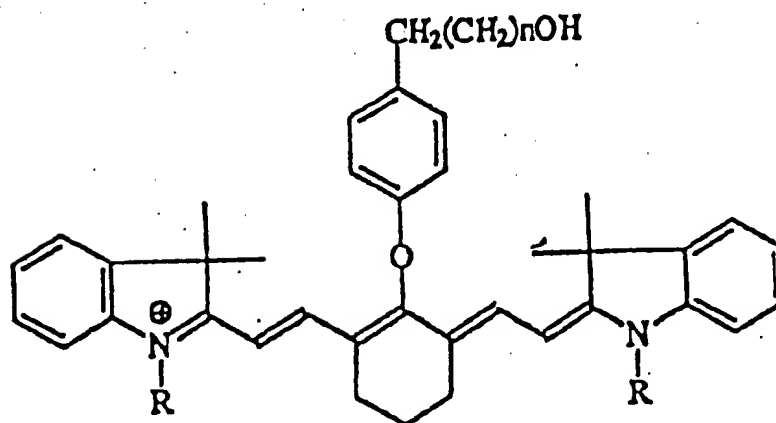
In the synthesized dyes shown by formula 5, an increase in quantum yield is obtained by coupling the nitrogen groups on the two sides of the heterocyclic base by a polymethylene group including a number of methylene monomers chosen to maintain the distance between them. Generally, the methylene chain is between four and ten groups long, with nine being preferable. Similarly, ethers may be used with the same number of groups. The ethers are more stable. A functional group may be attached to provide convenient coupling to biomolecules.

A proposed series of molecules represented by formula 6 is expected to have increased solubility to permit its use in an aqueous medium through the use of an amino group or ether group attached to the benzene ring. In this series, X may be oxygen or NH, Y may be NCS (isothiocyanate) or H, and R may be H, NCS, CH<sub>2</sub>OH, CH<sub>2</sub>NCS or COOH according to tables 1 or 2.

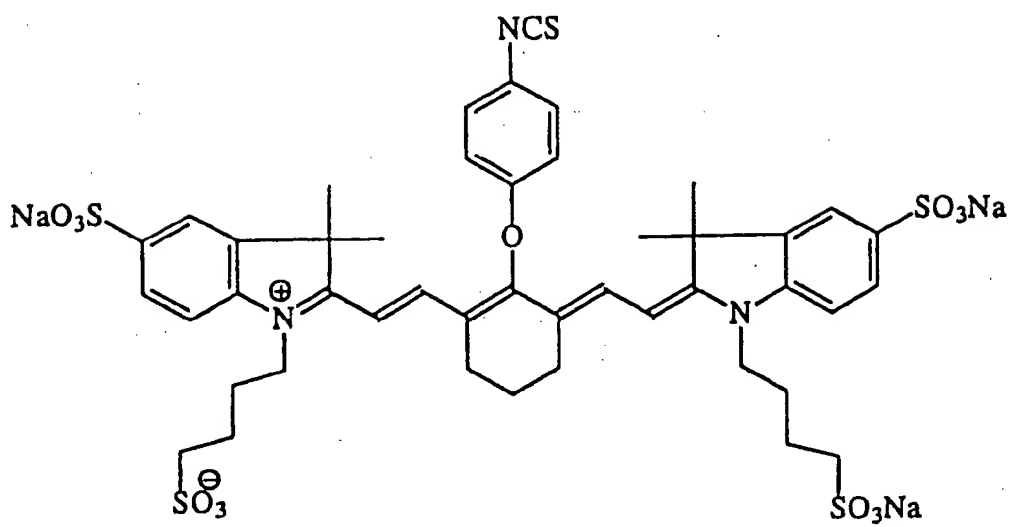
FORMULA 1



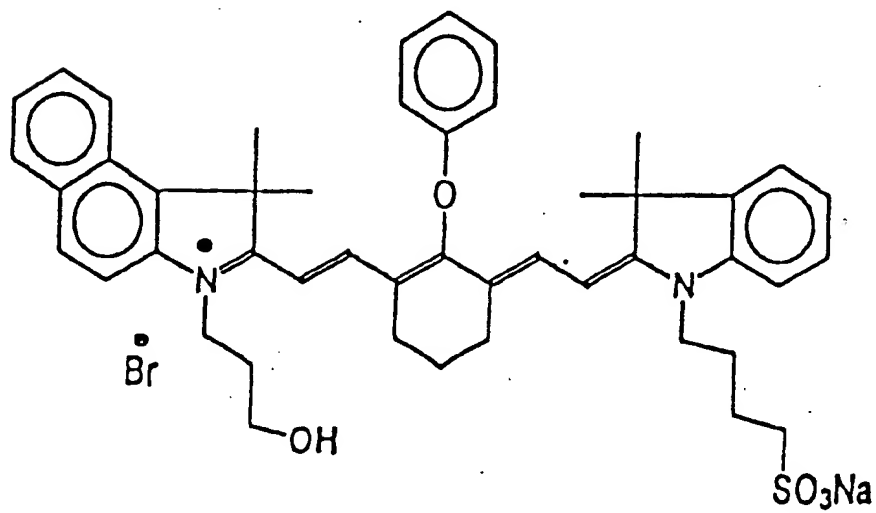
FORMULA 2



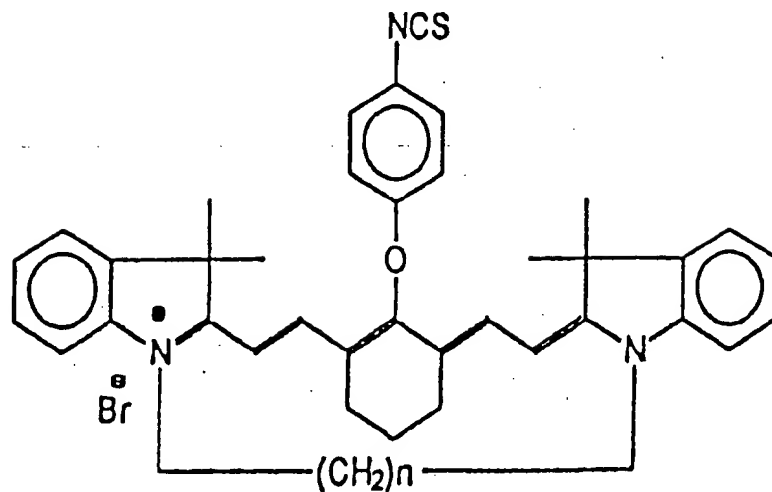
FORMULA 3



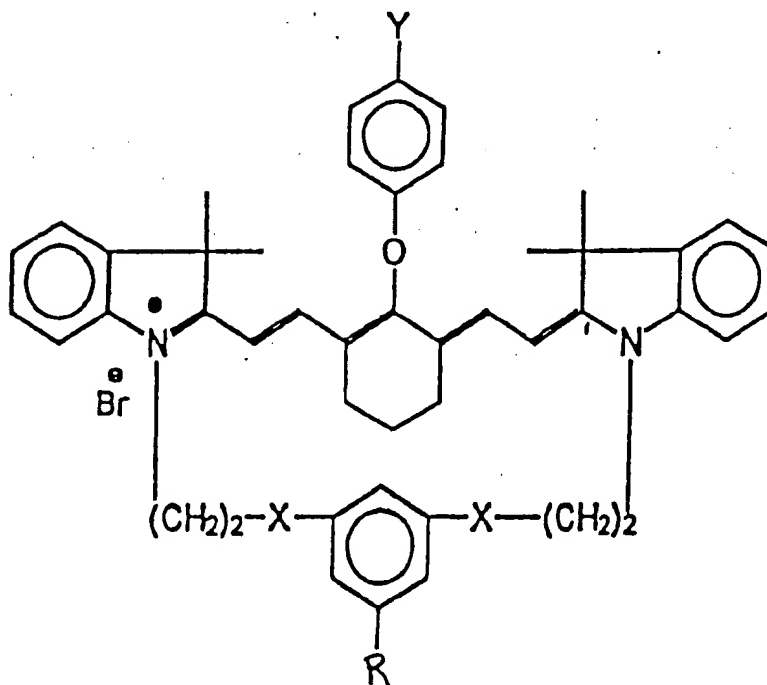
FORMULA 4



FORMULA 5



FORMULA 6



In formula 7, representing synthesized dyes:  $R_1$  may be ethyl, 4-sulfonatobutyl, 3-aminopropyl, phthalimidobutyl, hexyl and pentyl carboxylates; Z may be S, O,  $CMe_2$ ,  $R_2$  may be H,  $SO_3Na$ ,  $SO_3^+Et_3NH$ ,  $OCH_3$ ,  $NO_2$ ,  $CO_2Na$ , or  $CO_2^-Et_3NH$  and X may be H,  $N(CH_2CO_2CH_3)_2$ ,  $CH_2CH_2OH$ , NCS,  $CH_2CH_2CH_2OH$ , or  $-(CH_2)_n-OH$  where the number n is any number between and including 8-12.

One series of synthesized molecules is that set of formula 7 where  $R_1$  is 4-sulfonatobutyl and Z is  $CMe_2$ . In this

series of synthesized molecules:

- (1) When X and R<sub>2</sub> are hydrogen, the absorbance wavelength is 767 nanometers, the emission wavelength is 787 nanometers and a high quantum yield of 29% in methanol is obtained.
- (2) When X is equal to MeO and R<sub>2</sub> is hydrogen, the absorbance wavelength is about 768 in both water and methanol nanometers and the emission wavelength is 789 nm. The quantum yield in methanol is 35%.
- (3) When X is a nitro group and R<sub>2</sub> is hydrogen, the absorbance wavelength is between 770 and 774 nanometers in water and methanol and the emission wavelength is between 792 and 798 nm wavelengths in water and methanol, with a high quantum yield of 27% in methanol.
- (4) When X is NCS and R<sub>2</sub> is hydrogen (shown in formula 8), the absorbance wavelength is about 769 nanometers in both water and methanol and the emission wavelength is 788 nm in water and methanol with a high quantum yield of 38% in methanol.
- (5) When X is NCS and R<sub>2</sub> is OMe, the absorbance maximum wavelength is between 790 and 796 nm in water and methanol and the maximum emission wavelength is at 814 nm in water and methanol. The quantum yield is low.

The series of molecules represented by formula 9 have been synthesized and some are effective as IR labels when R<sub>1</sub>, R<sub>2</sub> and X are as shown in table 3.

The series of molecules represented by formula 10 have been synthesized and are effective as IR labels when R<sub>1</sub> and X are as shown in table 4.

The series of molecules represented by formula 11 have been synthesized and are effective as IR labels when R<sub>1</sub>, R<sub>2</sub> and X are as shown in table 5.

The series of molecules represented by formula 12 have been synthesized and are effective as IR labels when R<sub>1</sub>, R<sub>2</sub> and X are as shown in table 6.

The series of molecules represented by formula 13 have been synthesized and are effective as IR labels when R and X are as shown in table 7.

In another embodiment, the infrared dyes are selected to provide at least two dyes with each dye having its excitation and/or radiation spectra spaced sufficiently from each other dye so the fluorescence from each dye can be distinguished from each other dye either by the wavelength that excites it into fluorescence, or the wavelength at which it fluoresces or both, or by the fluorescence lifetime of the dye. The wavelength spacing is maintained sufficiently close to be excited by laser diodes. The dyes may be incorporated in probes and primers for attachment to oligonucleotides by such methods as described, for example, in Ruth, Jerry L. (1984) DNA 3, 123. The dyes may be newly synthesized or prepared by modifying existing, commercially available dyes.

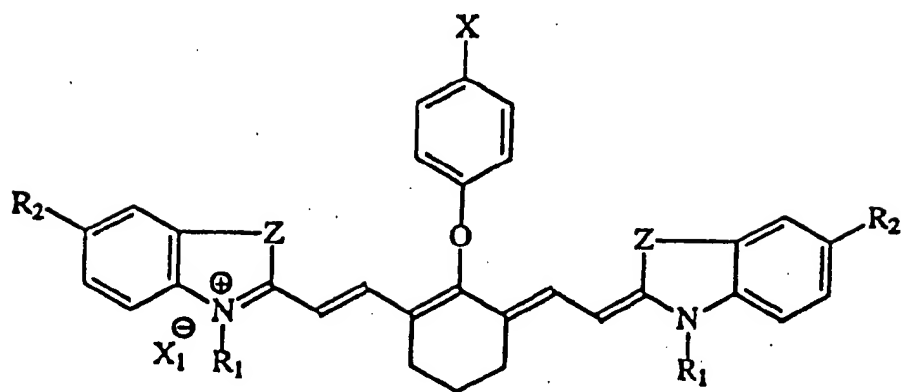
There are many dyes suitable for such modification such as for example: (1) 3,3'-Diethylthiadibocyanine Iodide; (2) 3,3'-Diethylthiatricbocyanine Perchlorate; (3) 3,3'-Diethyloxatricbocyanine Iodide; (4) 1,1',3,3,3'-Hexamethylindotricbocyanine Perchlorate; (5) 1,1'-Diethyl-2,2'-dicarbocyanine Iodide; (6) 3,3'-Diethylthiadibocyanine Iodide; (7) 3,3'-Diethyloxatricbocyanine Iodide; (8) 1,1',3,3,3,3'-Hexamethylindotricbocyanine Perchlorate; (9) 1,1',3,3,3,3'-Hexamethylindotricbocyanine Iodide; and (10) Indocyanine Green.

In one embodiment, the dye has the chemical structure shown in formula 14, with R being -CH<sub>2</sub>CH<sub>3</sub>. This dye is close to having the desired wavelength of maximum fluorescence and the wavelength of maximum absorbance may be modified by changing the functional group R. The unmodified dye may be obtained from Laboratory and Research Products Division, Eastman Kodak Company, Rochester, New York 14650. It is advertised in the Kodak laser dyes, Kodak publication JJ-169.

The modifications with 1,3-propanediol can be made in a manner known in the art as illustrated by equation 1. For example, changes occur when different esters are formed replacing the ethyl alcohol in the original dye molecule (R equal -CH<sub>2</sub>-CH<sub>3</sub> of formula 14). If different glycol esters are formed, absorbance maxima of these new near infrared dyes shift to the longer wavelengths.

The absorbance maximum is dependent on the distance of the oxygen atoms in the glycol functional group. However, the fluorescence maxima of these new near infrared dyes are practically at same wavelength of the dye of formula 14, i.e. 819 nm. This indicates that only the excitation process has changed, i.e. only the energy level to which the transition occurs during

## FORMULA 7



## FORMULA 8

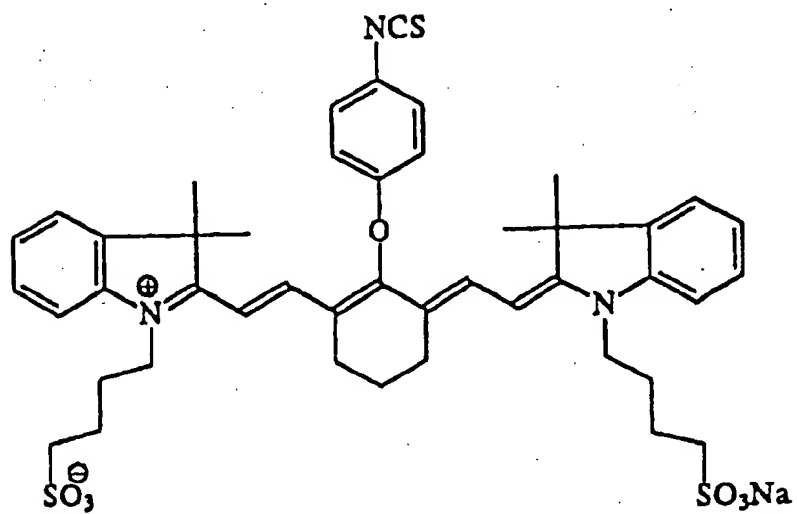


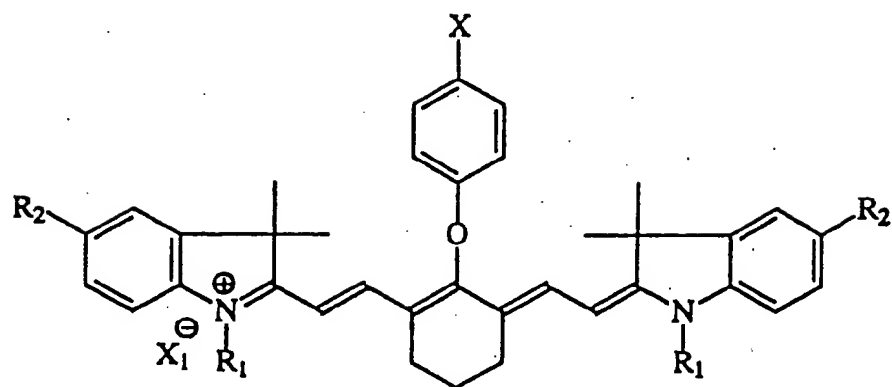
TABLE 1

Dye formula 6 series, when X = O (oxygen)		
X	Y	R
O	NCS	H
O	H	NCS
O	H	CH <sub>2</sub> OH
O	H	CH <sub>2</sub> NCS
O	H	COOH

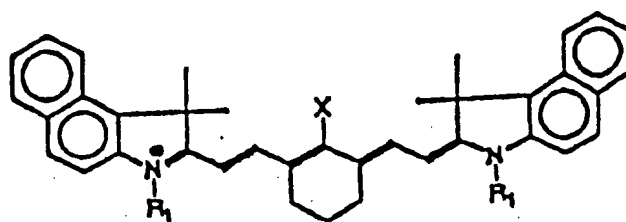
TABLE 2

Dye formula 6 series, when X = NH		
X	Y	R
NH	NCS	H
NH	H	NCS
NH	H	CH <sub>2</sub> OH
NH	H	CH <sub>2</sub> NCS
NH	H	COOH

## FORMULA 9



## FORMULA 10



## FORMULA 11

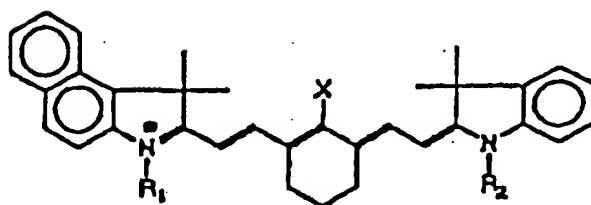


TABLE 3

R <sub>1</sub>	R <sub>2</sub>	X
-C <sub>4</sub> H <sub>8</sub> SO <sub>3</sub> <sup>-</sup> (Na <sup>+</sup> )	H	NCS
-C <sub>4</sub> H <sub>8</sub> SO <sub>3</sub> <sup>-</sup> (Na <sup>+</sup> )	H	OCH <sub>3</sub>
-C <sub>4</sub> H <sub>8</sub> SO <sub>3</sub> <sup>-</sup> (Na <sup>+</sup> )	H	NO <sub>2</sub>
-C <sub>4</sub> H <sub>8</sub> SO <sub>3</sub> <sup>-</sup> (Na <sup>+</sup> )	H	H
-C <sub>2</sub> H <sub>5</sub> (I <sup>-</sup> )	H	OCH <sub>3</sub>
-C <sub>2</sub> H <sub>5</sub> (I <sup>-</sup> )	OCH <sub>3</sub>	OCH <sub>3</sub>
-C <sub>4</sub> H <sub>8</sub> SO <sub>3</sub> <sup>-</sup> (Na <sup>+</sup> )	OCH <sub>3</sub>	OCH <sub>3</sub>
-C <sub>4</sub> H <sub>8</sub> SO <sub>3</sub> <sup>-</sup> (Na <sup>+</sup> )	OCH <sub>3</sub>	NCS
-C <sub>3</sub> H <sub>6</sub> NCS(Br <sup>-</sup> )	H	H
C <sub>2</sub> H <sub>5</sub> (I <sup>-</sup> )	H	NCS
C <sub>4</sub> H <sub>8</sub> SO <sub>3</sub> <sup>-</sup> (Na <sup>+</sup> )	SO <sub>3</sub> <sup>-</sup> Na <sup>+</sup>	NCS
C <sub>2</sub> H <sub>5</sub> (I <sup>-</sup> )	SO <sub>3</sub> <sup>-</sup> Na <sup>+</sup>	NCS

Table 4

R <sub>1</sub>	X
-C <sub>2</sub> H <sub>5</sub> (I <sup>-</sup> )	OPhNCS
-C <sub>4</sub> H <sub>8</sub> SO <sub>3</sub> <sup>-</sup> (Na <sup>+</sup> )	OPhNCS
-C <sub>3</sub> H <sub>6</sub> NCS(Br <sup>-</sup> )	OPh
phthalimido butyl	OPhNCS

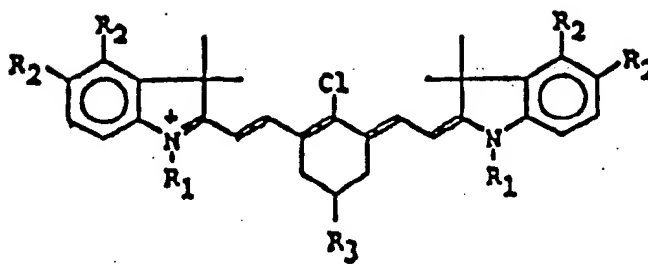
Table 5

R <sub>1</sub>	R <sub>2</sub>	X
-C <sub>3</sub> H <sub>6</sub> NCS(Br <sup>-</sup> )	-C <sub>4</sub> H <sub>8</sub> SO <sub>3</sub> H	Cl
-C <sub>3</sub> H <sub>6</sub> NCS(Br <sup>-</sup> )	-C <sub>2</sub> H <sub>5</sub>	Cl
-C <sub>3</sub> H <sub>6</sub> NCS	-C <sub>4</sub> H <sub>8</sub> SO <sub>3</sub> <sup>-</sup>	OPh
-C <sub>3</sub> H <sub>6</sub> NCS(Br <sup>-</sup> )	-C <sub>2</sub> H <sub>5</sub>	OPh
C <sub>3</sub> H <sub>6</sub> NCS	C <sub>4</sub> H <sub>8</sub> SO <sub>3</sub> <sup>-</sup>	OPh(pMeO)

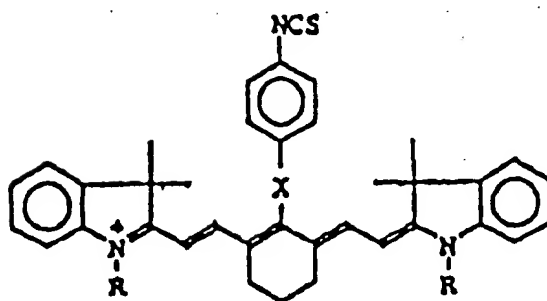
Table 6

R <sub>1</sub>	R <sub>3</sub>	R <sub>2</sub>
C <sub>2</sub> H <sub>5</sub> (I <sup>-</sup> )	H	H
(CH <sub>2</sub> ) <sub>4</sub> SO <sub>3</sub> <sup>-</sup> (Na <sup>+</sup> )	H	H
C <sub>2</sub> H <sub>5</sub> (I <sup>-</sup> )	COOCH <sub>3</sub>	H
C <sub>2</sub> H <sub>5</sub> (I <sup>-</sup> )	COOCH <sub>3</sub>	-(CH) <sub>4</sub> -
(CH <sub>2</sub> ) <sub>4</sub> SO <sub>3</sub> <sup>-</sup> (Na <sup>+</sup> )	COOCH <sub>3</sub>	-(CH) <sub>4</sub> -

## FORMULA 12



## FORMULA 13



## FORMULA 14

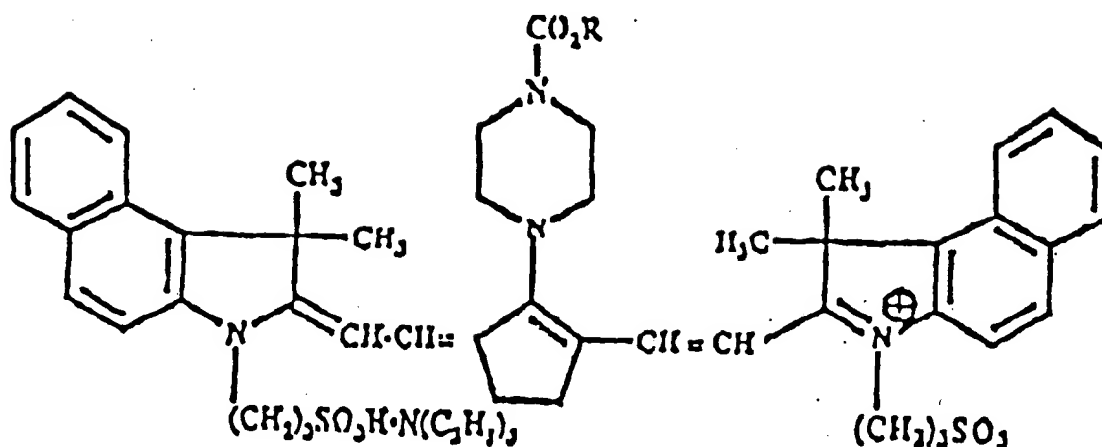


TABLE 7

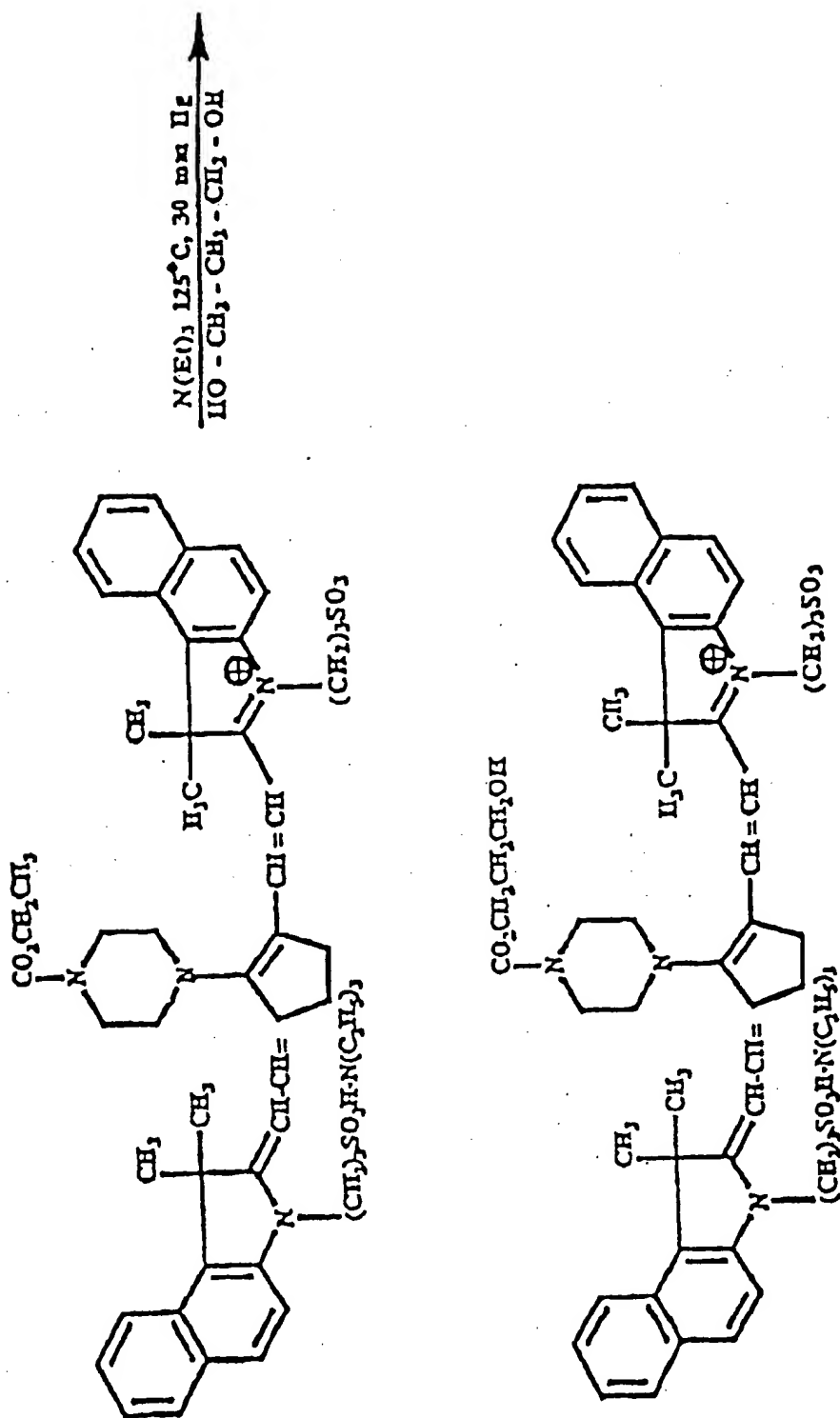
R	X
C <sub>2</sub> H <sub>5</sub> (I <sup>-</sup> )	O
(CH <sub>2</sub> ) <sub>4</sub> SO <sub>3</sub> <sup>-</sup> (Na <sup>+</sup> )	O
C <sub>2</sub> H <sub>5</sub> (I <sup>-</sup> )	S
(CH <sub>2</sub> ) <sub>4</sub> SO <sub>3</sub> <sup>-</sup> (Na <sup>+</sup> )	S

absorption has changed. The lowest vibronic level of first excited state remains unchanged. The absorbance maxima of such esters derived from several glycols are: (1) ethylene glycol 796 nm (nanometers); (2) 1,3-Propanediol 780 nm; (3) 1,4- Butanediol 754 nm; (4) 1,6-Hexanediol 745 nm; (5) Triethylene glycol 790 nm; (6) ethylene thiol 810 nm; and (7) IR-144 (R=CH<sub>2</sub>CH<sub>3</sub>) 742 nm.

In the preferred embodiment, the fluorescence maximum wavelength is about 819 nanometers and the detector is adjusted to receive this wavelength and not others by appropriate filtering. The absorbance maxima is selected to be different and to correspond to the preferred available laser diode emission. For example, in formula 14, R may be any of the following seven groups, depending on the desired wavelength of the emitted light, which are:

- (1) -CH<sub>2</sub>CH<sub>2</sub>-OH for an absorbance wavelength of 796 nanometers;
- (2) -CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>-OH for an absorbance wavelength of 780 nanometers;
- (3) -CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OH for an absorbance wavelength of 754 nanometers;
- (4) -CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>-OH for an absorbance wavelength of 745 nanometers;

EQUATION 1



- (5)  $-CH_2CH_2-O-CH_2CH_2-O-CH_2CH_2-OH$  for an absorbance wavelength of 790 nanometers;  
 (6)  $-CH_2CH_2-SH$  for an absorbance wavelength of 810 nanometers; and  
 (7)  $-CH_2CH_3$  for an absorbance wavelength of 742 nanometers.

More specifically, in the uncatalyzed synthesis of new heptamethine cyanine dyes which are potential precursors

for making functionalized near-infrared labels, a mixture of a quaternary salt of a heterocyclic base containing an activated methyl group (2 equiv) and 2-chloro-1-formyl-3-(hydroxymethylene)cyclohex-1-ene (2), an unsaturated bisaldehyde derived from cyclohexanone (1 equiv) is heated to reflux in a mixture of 1-butanol and benzene (7:3) as solvent, without using any catalyst. The water formed during the reaction is removed as an azeotrope by a Dean-Stark condenser. The reactions generally require 3 - 12 h for completion. The resulting product is generally pure after simple filtration of the dye from the crude reaction mixture followed by washings with diethyl ether. A wide range of 2-methyl-1-alkyl quaternary salts of various indole and benzindole derivatives undergo this reaction in a facile manner to form the corresponding symmetric dyes. An important feature of the current method is that the slower rate of the reaction allows one to prepare nonsymmetric dyes derived from two different heterocycles in a single pot in a fairly good yield. The syntheses of these nonsymmetric dyes become important when changes in the spectral and physical properties of the dyes are desired for specific application and compatibility with instrumentation. These changes can be incorporated by appropriately modifying the structural design of the dyes.

Dye synthesis is illustrated by chart 1, which shows the synthesis of formulae 2, 3, 7, 8, and 9. In the chloro intermediate in chart 1,  $R_1$  and  $R_2$  may have the values shown in table 9. In chart 1,  $X_1$  is a halogen. Chart 2 shows the synthesis of formula 10. Chart 3 shows the synthesis of formulae 1, 4 and 11. Chart 4 shows the synthesis of formula 12. Chart 5 shows the synthesis of formula 13. In chart 5, X is oxygen or sulfur.

In FIG. 1, there is shown a perspective view of an embodiment of sequencer in which the method of this invention may be performed. This sequencer is described in structure and operation in the aforementioned United States patent application 07/570,503 filed August 21, 1990; United States patent application 07/078,279 filed July 27, 1987; and United States Patent 4,729,947, all

Table 9

$R_1$	$R_2$
$C_4H_8SO_3H$	H
$C_2H_5$	H
$C_3H_5NH_3Br^-$	H
$C_4H_8SO_3H$	$SO_3^-Na^+$
$C_2H_5$	$SO_3^-Na^+$
$C_4H_8SO_3H$	$OCH_3$
$C_2H_5$	$OCH_3$

CHART 1

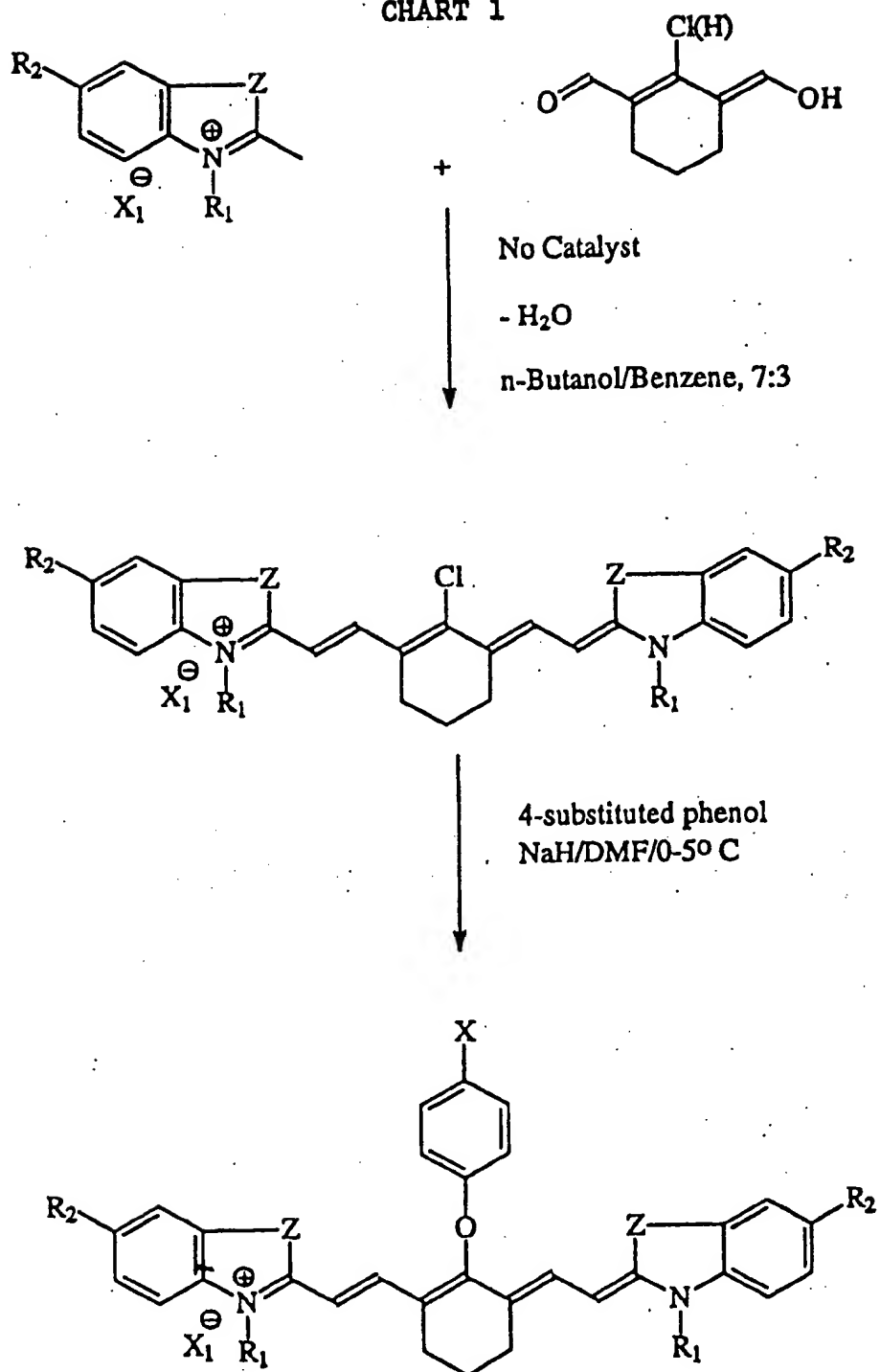


CHART 2

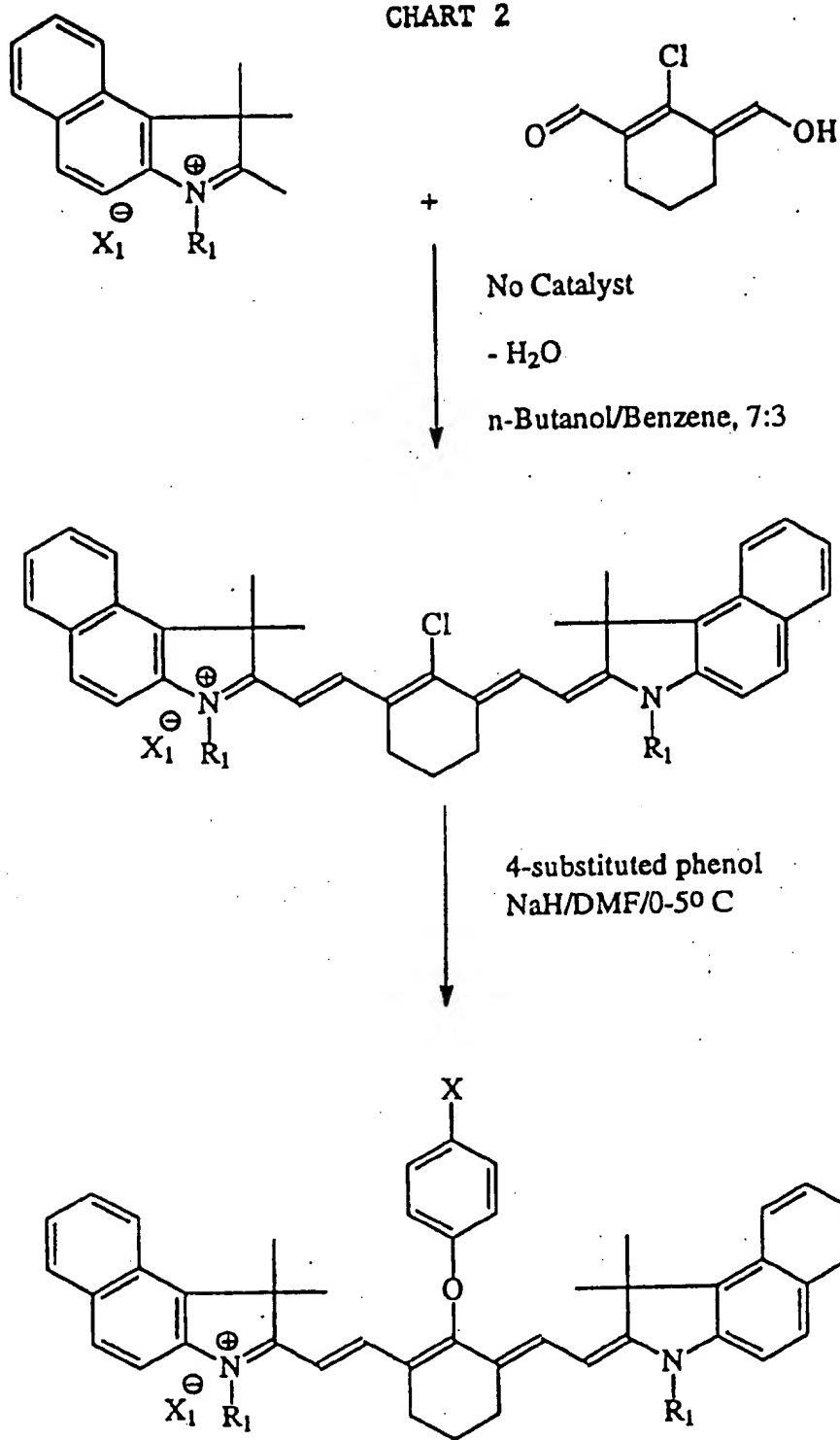


CHART 3

5

10

15

20

25

30

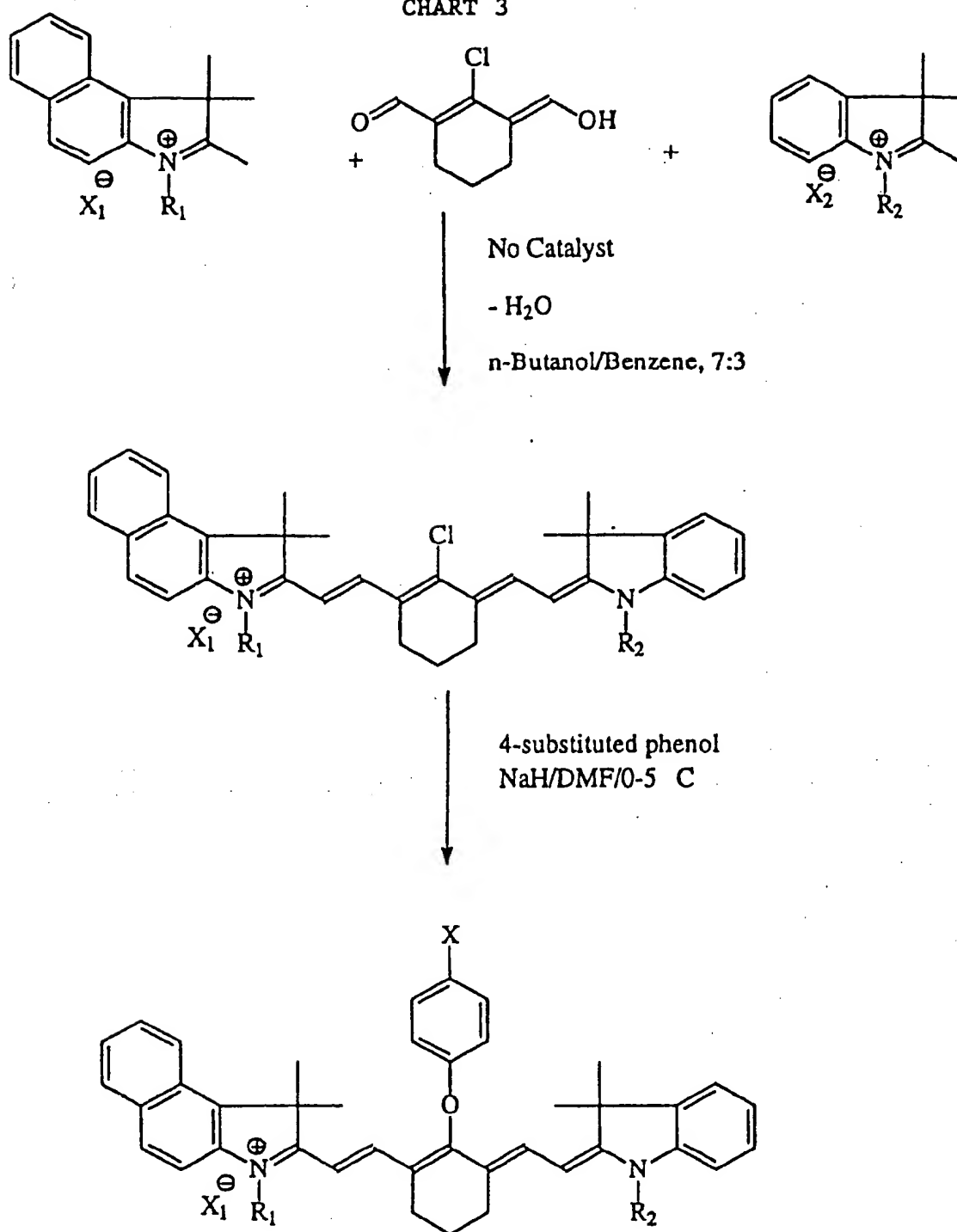
35

40

45

50

55



## CHART 4

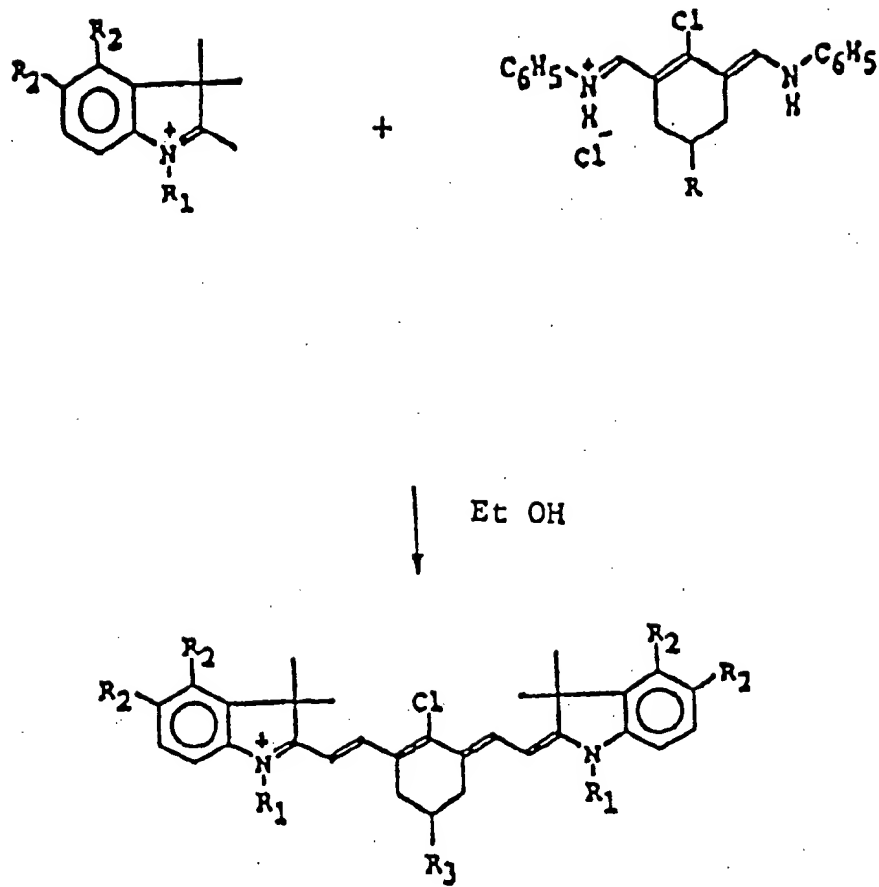
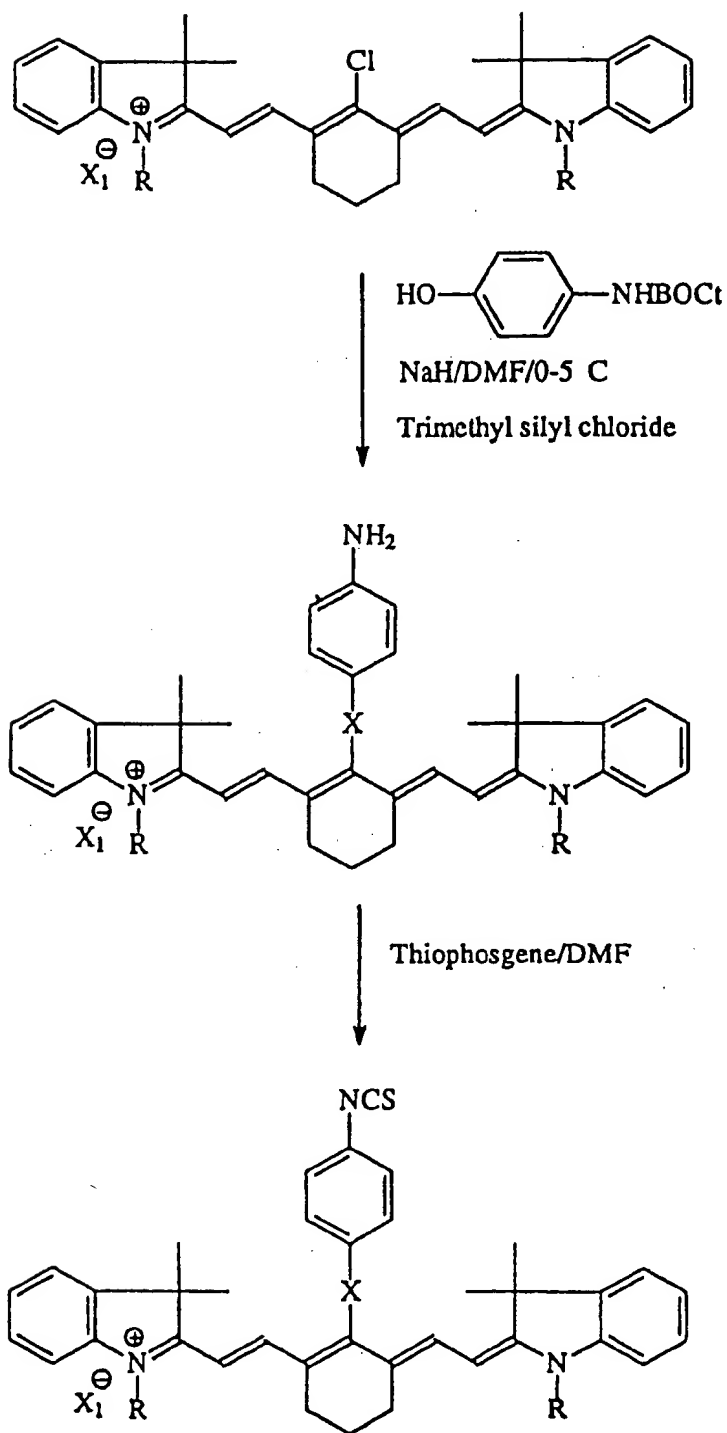


CHART 5



of which are entitled DNA SEQUENCING and which were filed by Middendorf et al.

In FIG. 2, there is shown a sectional view of the remote station 122A taken through section lines 2-2 of FIG. 1 having an electrophoresis section 140, a scanning section 142, an electrophoresis power supply 144, a system power supply section 144A, an analog board 146 and a digital board 148. The electrophoresis section 140 is positioned near the

front of the cabinet and a portion of it is adapted to be scanned by the scanning section 142 in cooperation with circuitry on the analog board 146 and the digital board 148. All of the apparatus are electrically connected to the power supply section 144A for such operation.

To separate different DNA fragments into bands, the electrophoresis section 140 includes a gel sandwich 150, an upper buffer assembly 152, support assembly 154, and a lower buffer assembly 151 positioned to enclose the bottom of the gel sandwich 150. In the embodiment of FIG. 2, the gel sandwich 150 is held vertically and its temperature is controlled during operation. Bands are separated by applying voltage to the upper buffer assembly 152 and lower buffer assembly 151 and scanned by the scanning section 142.

To support the gel sandwich 150, the support assembly 154 includes a pair of upper side brackets and lower side brackets 160 and 162 (only one of each pair being shown in FIG. 2), an apparatus support plate 168, a temperature control heating plate 164, and a plastic spacer, shown at 166A-166C, in FIG. 2. The entire structure is supported on the apparatus support plate 168 which mounts the upper and lower side brackets 160 and 162.

The upper and lower side brackets 160 and 162 are each shaped to receive a pin (not shown) extending from one or the other side of the gel sandwich 150 and hold it in place in juxtaposition with the scanning section 142. The spacers as shown as 166A-166C space the temperature control heating plate 164 from the apparatus support plate 168 and maintain it at a constant selected temperature above ambient temperature. In the preferred embodiment, the temperature is maintained at 45-50 degrees Centigrade and should be maintained in a range of 30 degrees to 80 degrees.

The scanning section 142 includes a laser diode assembly (not shown in FIG. 2), a microscope assembly 172, a photodiode section 174 and a scanner mounting section 176. The laser diode assembly (not shown in FIG. 2) is positioned at an angle to an opening in the heating plate 164 so that light impinges on the gel sandwich 150 to cause fluorescence with minimum reflection back through the microscope assembly 172.

To receive the fluorescent light, the microscope assembly 172 is focused on the gel sandwich 150 and transmits fluorescent light emitted therefrom into the photodiode section 174 which converts it to electrical signals for transmission to and processing by the analog and digital boards 146 and 148 which may provide further analysis of data. The scanning section 142 moves along a slot in the apparatus support plate 168 which is mounted to the scanner mounting section 176 during this operation in order to scan across the columns in the gel sandwich 150.

The scanner mounting section 176 includes a mounting plate 180, a bearing plate 182, a stepping motor 184, a slidable support 186 and a belt and pulley arrangement 185, 188A and 188B. The mounting plate 180 is movably mounted to the apparatus support plate 168 through a frame member and supports the elongated bearing plate 182, the stepping motor 184 and two pulleys 188A and 188B. The elongated bearing plate 182 extends the length of the gel sandwich 150.

To permit motion of the laser diode assembly (not shown) and microscope assembly 172 with respect to the gel sandwich 150, the slidable support 186 supports the microscope assembly 172 and diode assembly and slidably rests upon the bearing plate 182. An output shaft 183 of the stepping motor 184 drives a pulley 188B through pulley 188, belt 185, and pulley 188A and the pulley 188B drives a belt (not shown) that is clamped to the slidable support 186 to move it the length of the gel sandwich 150 during scanning by the laser diode and microscope assembly 172 which rest upon it. The stepping motor 184 under the control of circuitry in the digital board 148 moves the pulley 188B to move the belt (not shown) and thus cause scanning across the gel sandwich 150.

As shown in this view, the electrophoresis power supply 144 is electrically connected to buffer in the upper buffer assembly 152 through an electrical connector 194 and to the lower buffer assembly 151 through a connector not shown in FIG. 2.

The upper buffer assembly 152 includes walls 197 forming a container to hold a buffer solution 195 and a cover 199 formed with a lip to fit over the walls 197 from the top and containing a downwardly extending flat member spaced away from the side walls and holding a conductor 211. The conductor 211 is electrically connected to the source of power through connector 194 which is mounted to the top of the cover 199 to permit electrical energization of the buffer solution 195.

The bottom buffer assembly 151 includes enclosed walls 201 defining a container for holding a buffer solution 203 and a cover 205 closing the container 201 and having a downwardly extending portion 213 extending into the buffer 203 for supporting a conductor 207 for applying energy to the bottom buffer solution 203. The gel sandwich 150 extends downwardly into the buffer solution 203 and upwardly into the buffer solution 195 to permit the electrical contact for electrophoresis. An "o" ring 197B provides a seal for the upper buffer assembly 152 so that the buffer solution 195 does not empty out of the upper buffer assembly 152.

In FIG. 3, there is shown a sectional view taken through lines 3-3 of FIG. 1 showing a portion of the electrophoresis section 140, a portion of the scanning section 142 (indicated twice in FIG. 3 for clarity) and the electrophoresis power supply section 144 (FIG. 2 only) mounted together to illustrate, from a top view, the arrangement of the apparatus support plate 168, the heater plate 164, the gel sandwich 150, a laser diode assembly 170, a microscope assembly 172 and a photodiode assembly 174. The heater plate 164 and apparatus support plate 168 have slots running in a horizontal direction orthogonal to the lanes of DNA in the electrophoresis section 140 sized to receive the ends of the laser diode assembly 170 and the microscope section 172 for scanning thereof.

To cooperate with the separation and scanning of DNA bands, the gel sandwich 150 includes a front glass plate 200, a gel section 202 and a rear glass plate 204 mounted in contact with the heater plate 164 and having a section exposed for scanning by the laser diode assembly 170 and the microscope assembly 172. The rear glass plate 204 contacts the heater plate 164 and is separated from the front glass plate 200 by the gel section 202 within which DNA separation takes place. The front and rear glass plates 200 and 204 may be any type of glass but are preferably soda lime which has low fluorescence in the far red and near infrared regions and is prepared by a process that provides optically flat surfaces without grinding.

To transmit light to the gel sandwich 150, the laser diode assembly 170 includes a housing 210, a focusing lens 212, a narrow band pass filter 214, a collimating lens 216 and a laser diode 218. The laser diode 218 emits far red, near infrared, or infrared light which is collimated by the laser collimating lens 216 and filtered through the narrow band pass filter 214. This light is focused by the focusing lens 212 onto the gel sandwich 150. Preferably, the point of focus on the gel section 202 of the gel sandwich 150 lies along or near the central longitudinal axis of the microscope section 172 and the photodiode section 174.

The thickness of the glass plates and the gel, the position of the laser and microscope assembly 172 and thus the angle of incidence and angle of reflection of the light from the laser and to the microscope assembly 172 are chosen, taking into consideration the refractive index of the gel and glass and the thickness of the glass plates and the gel, so that the light from the laser is maximally transmitted to the gel. The light from the laser is not directly reflected back because the angle of incidence to normal is equal to the Brewster's angle at the first interface and is such as to impinge on the markers with full intensity after refraction but not be reflected by the first surface of the gel sandwich 150 into the microscope assembly 172 and the microscope assembly 172 views those markers that fluoresce in its line of sight.

To maintain temperature control over the laser diode, the housing 210: (a) is coupled to a heat sink through a thermal electric cooler 220, and (b) encloses the focusing lens 212, narrow band pass filter 214, collimating lens 216 and laser diode 218; and (c) accommodates the electrical leads for the diode.

To receive and focus light emitted by fluorescent markers from the gel section 202 in response to the light from the laser diode assembly 170, the microscope assembly 172 includes a collection lens 230, a housing 232 and a focusing motor. The microscope assembly 172 is adapted to be positioned with its longitudinal axis centered on the collection lens 230 and aligned with the photodiode section 174 to which it is connected. For this purpose, the housing 232 includes a central passageway in which are located one or more optical filters (not shown) with a pass band matching the emission fluorescence of the marked DNA strands. With this arrangement, the collection lens 230 receives light from the fluorescent material within the gel section 202 and collimates the collected light for optical filtering and then transmission to the photodiode assembly 174.

To generate electrical signals representing the detected fluorescence, the photodiode assembly 174 includes a housing 240 having within it, as the principal elements of the light sensors, an inlet window 242, a focusing lens 244, a sapphire window 246 and an avalanche photodiode 248. To support the avalanche photodiode 248, a detector mounting plate 250 is mounted within the housing 240 to support a plate upon which the avalanche photodiode 248 is mounted. The inlet window 242 fits within the housing 240 to receive light along the longitudinal axis of the photodiode assembly 174 from the microscope assembly 172.

Within the housing 240 of the photodiode assembly 174, the sapphire window 246 and avalanche photodiode 248 are aligned along the common axis of the microscope assembly 172 and the photodiode assembly 174. The focusing lens 244 focuses light transmitted by the microscope assembly 172 onto a small spot on the avalanche photodiode 248 for conversion to electrical signals. A thermoelectric cooler 252 utilizing the Peltier effect is mounted adjacent to the detector mounting plate 250 to maintain a relatively cool temperature suitable for proper operation of the avalanche photodiode 248.

As best shown in this view, the stepping motor 184 rotates the belt 185 to turn the pulley 188A, which, in turn, rotates pulley 188B. The pulley 188B includes a belt 177 extending between it and an idler pulley 179 and attached at one location to the slideable support 186 (shown only in FIG. 2) to move the scanning microscope and laser lengthwise along the gel sandwich 150 for scanning purposes. The motor 184, by moving the carriage back and forth, accomplishes scanning of the gel sandwich 150.

In FIG. 4, there is shown a fragmentary perspective view of the gel sandwich 150 and the upper buffer assembly 152 mounted to each other showing the outer glass plate 200 cut away from the rear glass plate 204 to expose the gel section 202 to buffer solution within the upper buffer assembly 152. With this arrangement, DNA samples may be pipetted between the glass plates 200 and 204 and moved downwardly by electrophoresis beyond the upper buffer assembly 152 and through the gel sandwich 150 to the bottom buffer (not shown in FIG. 4).

In FIG. 5, there is shown a broken away view of the gel sandwich 150 illustrating the upper buffer assembly 152 and the lower buffer assembly 151 connected to it at each end. As shown in this view, the cover 199 includes a connecting post 214 which receives the conductor 211 for connection to the downwardly extending portion of the cover 199 into the buffer compartment. Between the glass plates 200 and 204 (FIG. 4) of the gel sandwich 150, are a plurality of downwardly extending recesses 231 in the gel section 202 (FIG. 4) between the plates. DNA sample is pipetted into these recesses to form channels for electrophoresing to the lower buffer assembly 151.

To form an electrical connection through the gel sandwich 150 from the upper buffer assembly 152 to the lower buffer assembly 151, a connecting post 218 is connected to the cover 205 of the lower buffer assembly 151 for receiving the conductor 207 which extends downwardly to the downwardly extended plate 213 and into the buffer solution.

In FIG. 6, there is shown a block diagram of the circuitry used to control the remote station 122A of the embodiment of FIG. 2 having a control, correlation and readout section 250, the scanner drive 176, the motor assembly 184 for moving the scanner drive 176, and the sensing configuration 252. The sensing configuration 252 includes the laser assembly 170 and the sensor assembly 174 which receives signals, removes some noise, and transmits the signals for display and readout in the control, correlation and readout section 250, while the scanner drive 176 and motor for the scanner drive 184 receive signals from the control, correlation and readout section 250 to control the motion of the sensor back and forth across the gel sandwich. This overall configuration is not part of the invention of this application except insofar as it cooperates with the sensing configuration 252 to scan the DNA and determine its sequence in accordance with the embodiments of FIGS 1-5.

To drive the sensor 174 from position to position, the motor assembly 184 includes a stepper motor 254 and a motor driver 256. The motor driver 256 receives signals from the control correlation and readout section 250 and actuates the stepper motor 254 to drive the scanner drive 176. The scanner drive 176 is mechanically coupled to a stepping motor 254 through a belt and pulley arrangement for movement back and forth to sense the electrophoresis channels on the gel sandwich 150 (FIG. 3). The stepping motor 254 and driver circuitry 256 are conventional and not themselves part of the invention.

The control, correlation and readout system 250 includes a computer which may be any standard microprocessor 260, a television display or cathode ray tube display 262 and a printer 264 for displaying and printing the results of the scans.

To sense data, the sensing configuration 252 includes, in addition to the laser 170 and the sensor 174, a chopper circuit 270, a sensor power supply 272, a preamplifier 274, a lock-in amplifier 276, a 6-pole filter 278, a 12-bit analogue to digital converter interface circuit 280 and a laser power supply 282.

The sensor 174 receives light from the laser 170 after it impinges upon the gel sandwich 150 (FIG. 3) and transmits the signals through preamplifier 274 to the lock-in amplifier 276. The sensor receives signals from the sensor power supply 272. The chopper circuit 270 provides pulses at synchronized frequencies to the lock-in amplifier 276.

The laser 170 receives power from the power supply 282 which is controlled by the chopper circuit 270 so that the signal from the laser is in synchronism with the signal applied to the lock-in amplifier 276 so that the output from the lock-in amplifier 276 to the 6-pole filter 278 discriminates against unwanted signal frequencies. This signal is converted to a digital signal in the 12-bit analogue to digital converter 280 which serves as an interface to the computer 260.

With this arrangement, the scanning rate may be set to discriminate against noise and the synchronized demodulation from the chopper control further reduces noise, particularly discriminating against the natural fluorescence of the glass in the gel sandwich 150 (FIGS. 2 and 3).

In FIG. 7, there is shown another embodiment C26A which may be substituted for the gel sandwich 150 and includes capillary columns 61, 63 and 65 as commonly used in capillary electrophoresis. These columns may be filled with buffer solution or a gel and be used for electrophoresis. In such a case, several capillaries may be used as a substitute for the gel sandwich 150 of the embodiment of FIGS. 2-5. Thus, the same band of A, G, C or T type bases might flow through several parallel bundles of capillaries, or they might flow through only one capillary per type of base; or bands of A, G, C, and T type bases may be combined to flow through only one capillary.

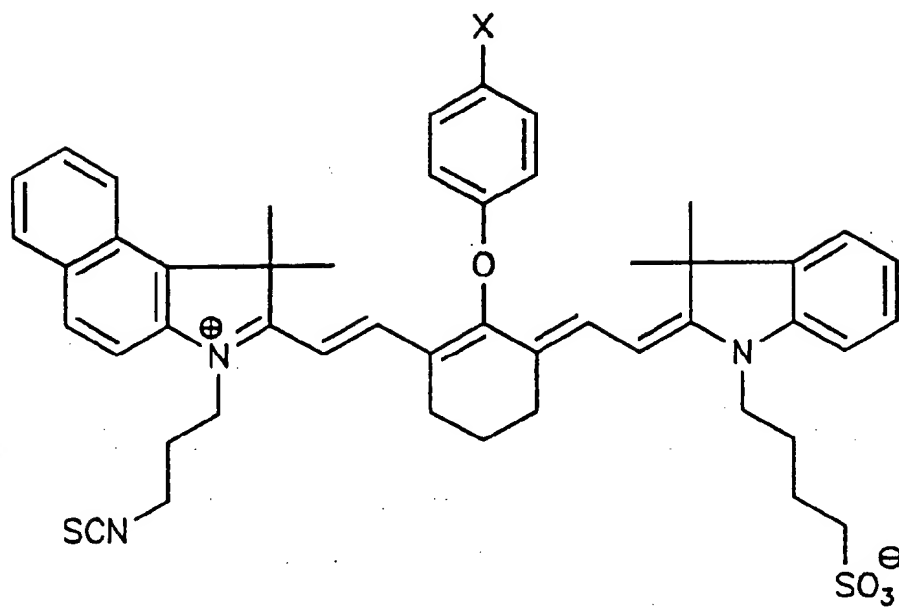
The separation path such as gel channels or capillary tube length should be no longer than two meters for a range of lengths of DNA from 50 to 10,000 or more bases. However, as the range of DNA lengths increase, the time required increases. Also, the time required for each separation is in the range of from 1/2 second to 5 minutes for each added base of length separation.

From the above summary, it can be understood that the techniques for the sequencing of fluorescence labeled DNA of this invention have several advantages, such as: (1) because the dyes have their emission spectra in the far red, near infrared or infrared light spectrum, small inexpensive diode lasers may be used; and (2) this wavelength region is characterized by relatively low background fluorescence in glass, and, therefore, less noise in the received signal.

Although a preferred embodiment of the invention has been described with some particularity, many modifications and variations are possible in the preferred embodiment within the light of the above description. Accordingly, within the scope of the appended claims, the invention may be practiced other than as specifically described.

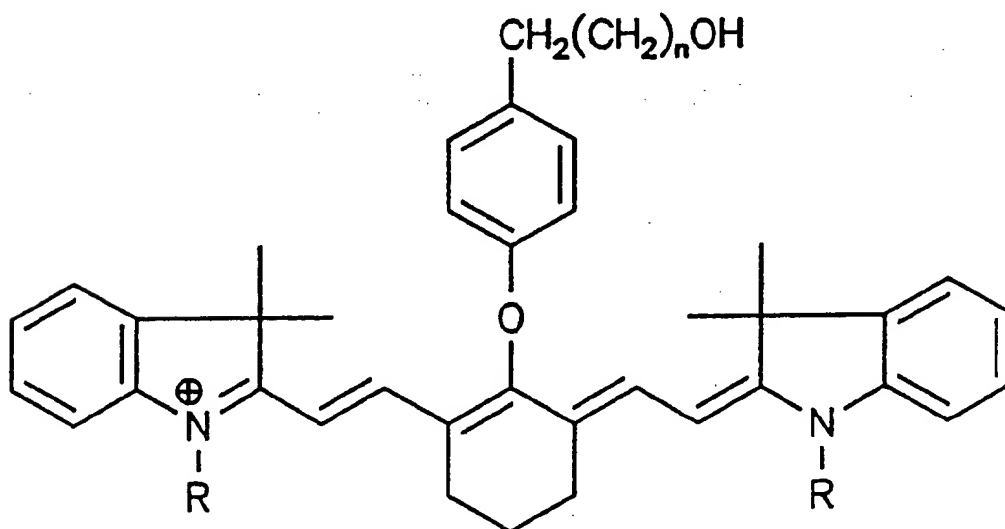
## Claims

1. A method of identifying strands of DNA comprising the steps of marking the strands with fluorescent labels that emit light irradiating the strands with light, and detecting the light emitted from the fluorescent labels, characterized in that the fluorescent labels emit light in a region of wavelengths including at least one wavelength within one of a far red, near infrared, and infrared regions wherein the fluorescent label includes a chromophore having the formula:



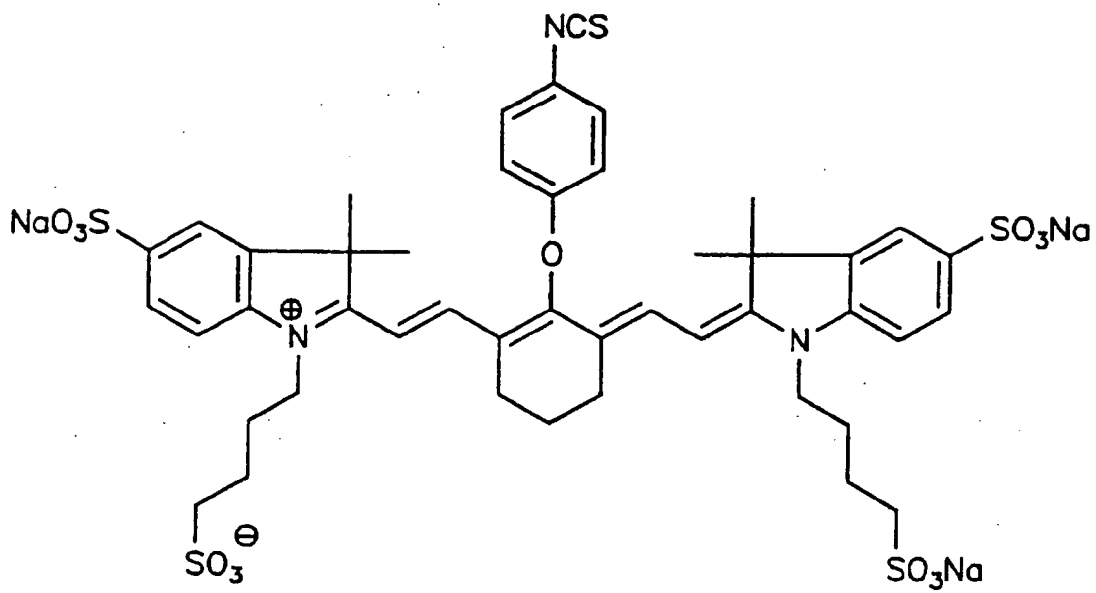
25

Where X is H or  $\text{OCH}_3$ ;  
or the formula:

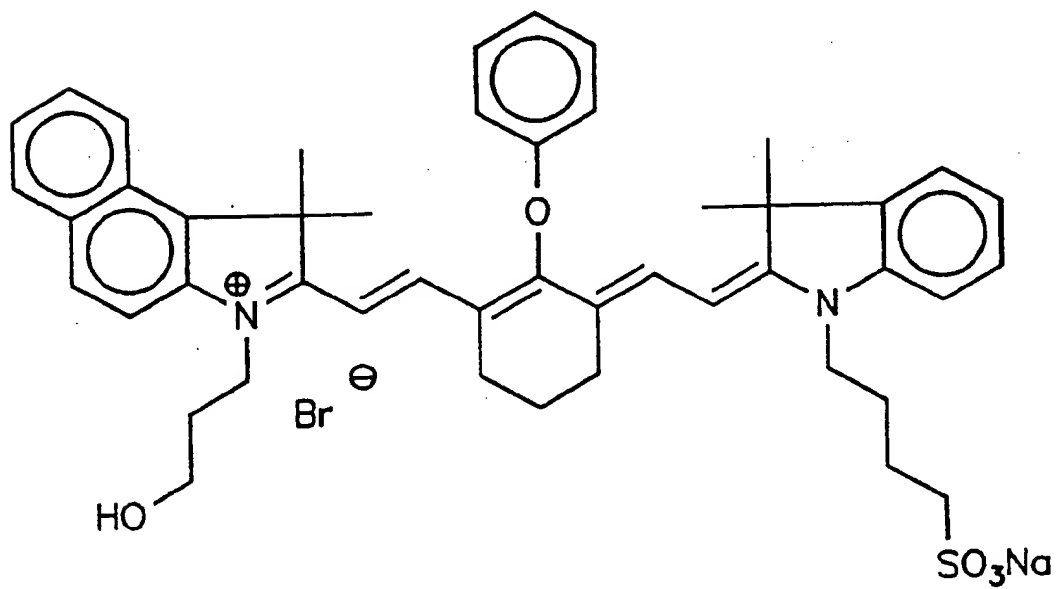


55

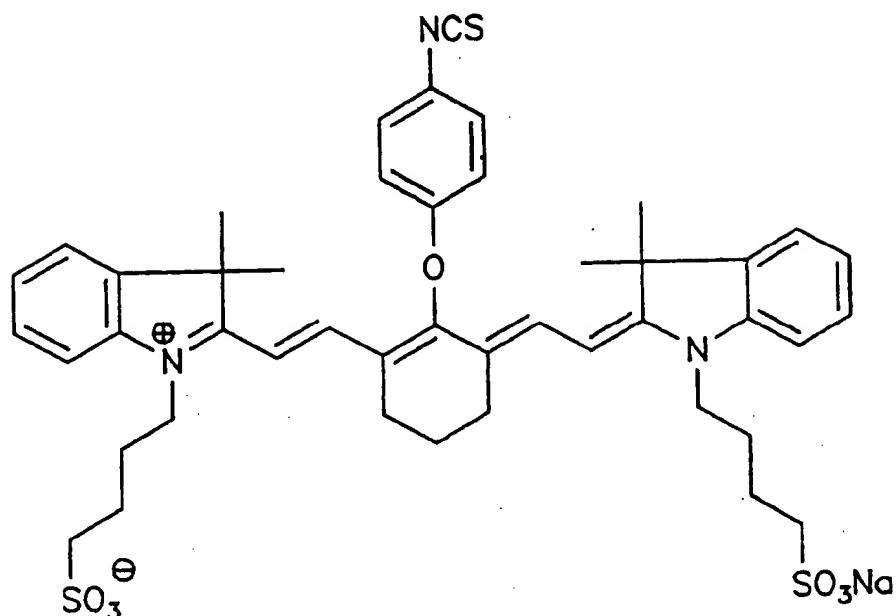
where R is ethyl or 4-sulfonatobutyl and n is either 1 or 2.  
or the formula



25 or the formula

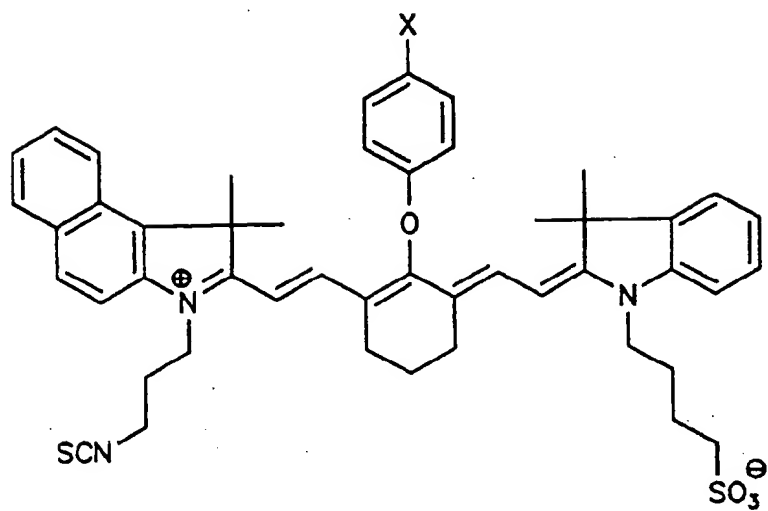


50 or the formula

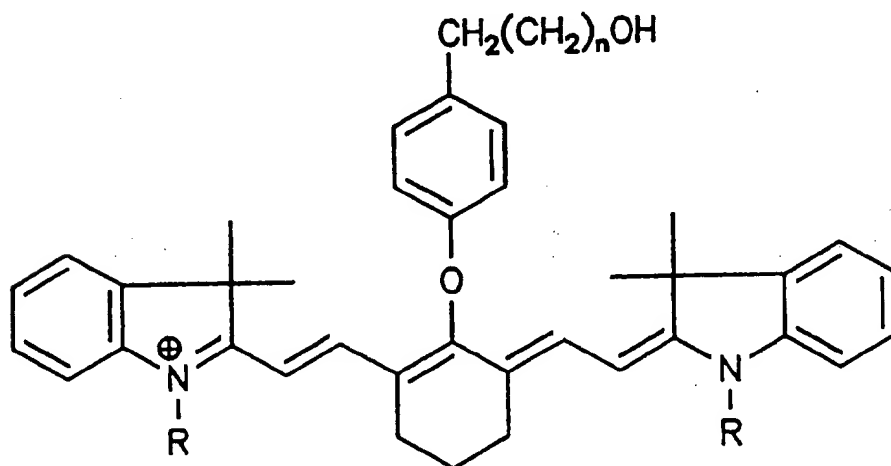


and the irradiating light has a wavelength within one of the far red, near infrared, and infrared regions.

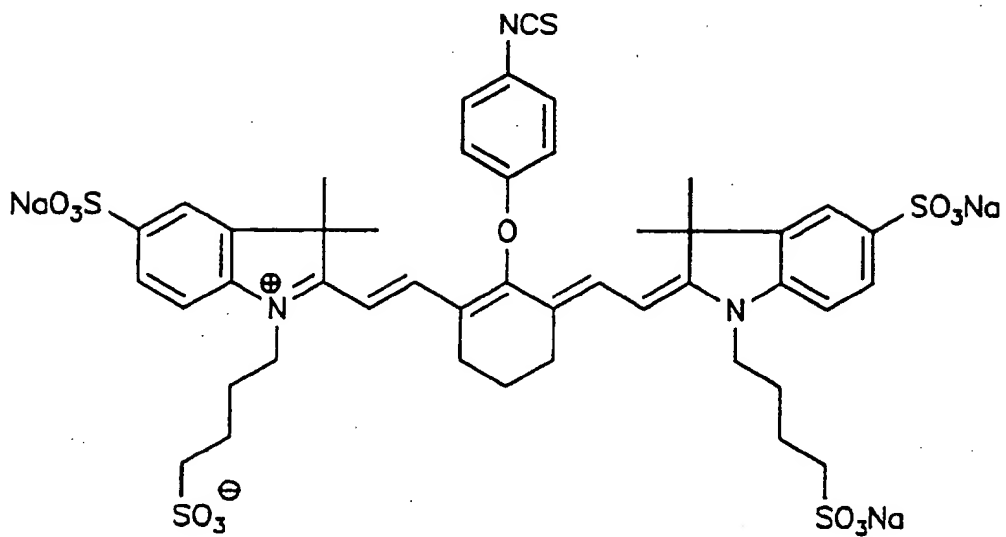
2. A method of identifying strands of DNA in accordance with claim 1 characterized in that the irradiating is with one of the far red, near infrared, and infrared laser diode light sources (170).
3. A method according to either claim 1 or claim 2 characterized by applying DNA samples marked with fluorescent dye having absorbance and fluorescence maxima at one of the far red, near infrared, and infrared wavelengths when combined with the DNA at a plurality of locations in a gel electrophoresis slab (150) or set of capillaries (61, 63, 65) for electrophoresing in a plurality of channels through the gel electrophoresis slab (150) or set of capillaries (61, 63, 65); establishing electrical potential across said gel electrophoresis slab (150) or set of capillaries (61, 63, 65) wherein DNA samples are resolved in accordance with the size of DNA fragments in said gel electrophoresis slab (150) or set of capillaries (61, 63, 65) into fluorescently marked DNA bands; and irradiating the separated samples photo-electrically while they are in the slab (150) or set of capillaries (61, 63, 65) with a laser diode (170) and a sensor (174) wherein the laser (170) scans with far red, near infrared, or infrared scanning light at the wavelength within an absorbance spectrum of said fluorescently marked DNA samples and sensing light at the emission wavelength of the marked DNA.
4. A dye characterized by one of the formula:



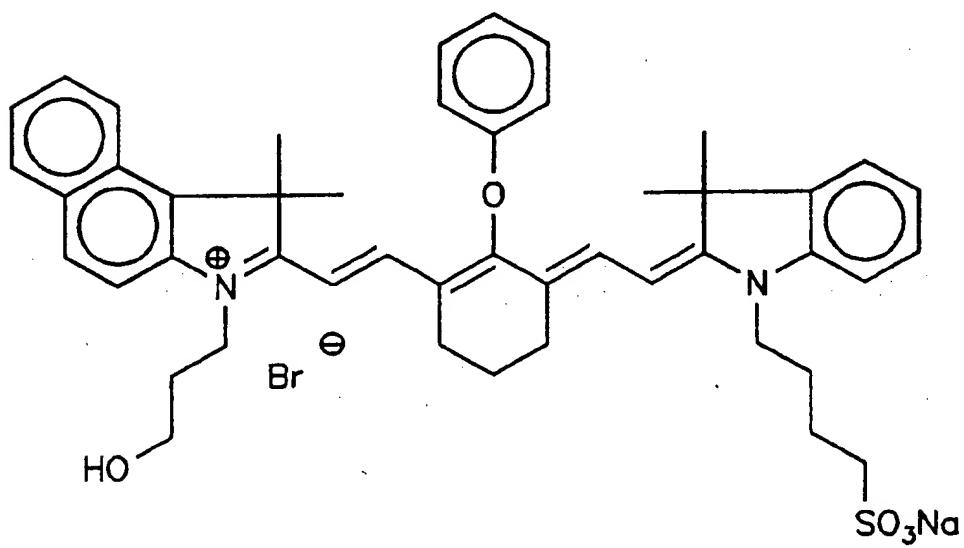
where X is H or OCH<sub>3</sub>  
or



where R is ethyl or sulfonatobutyl and n is either 1 or 2;  
or



or



or

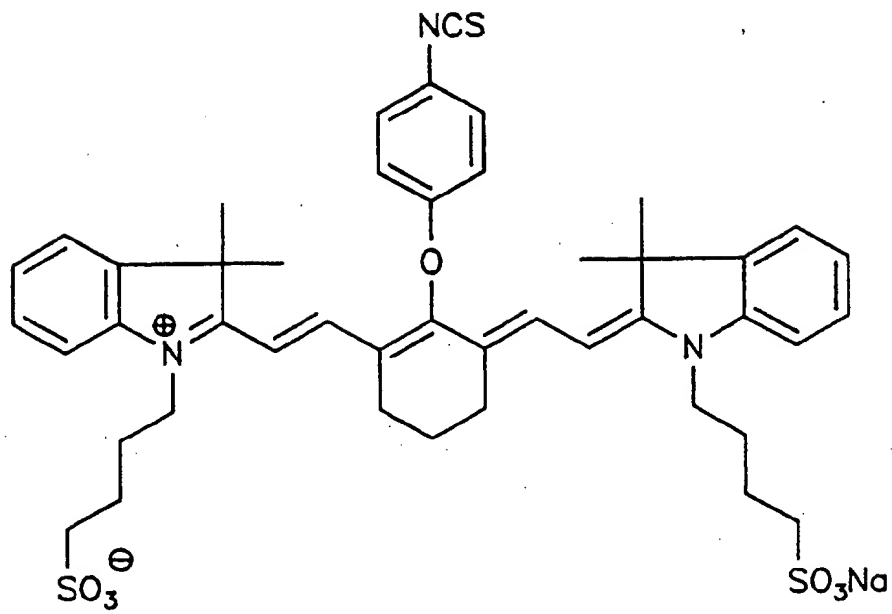


FIG. 1

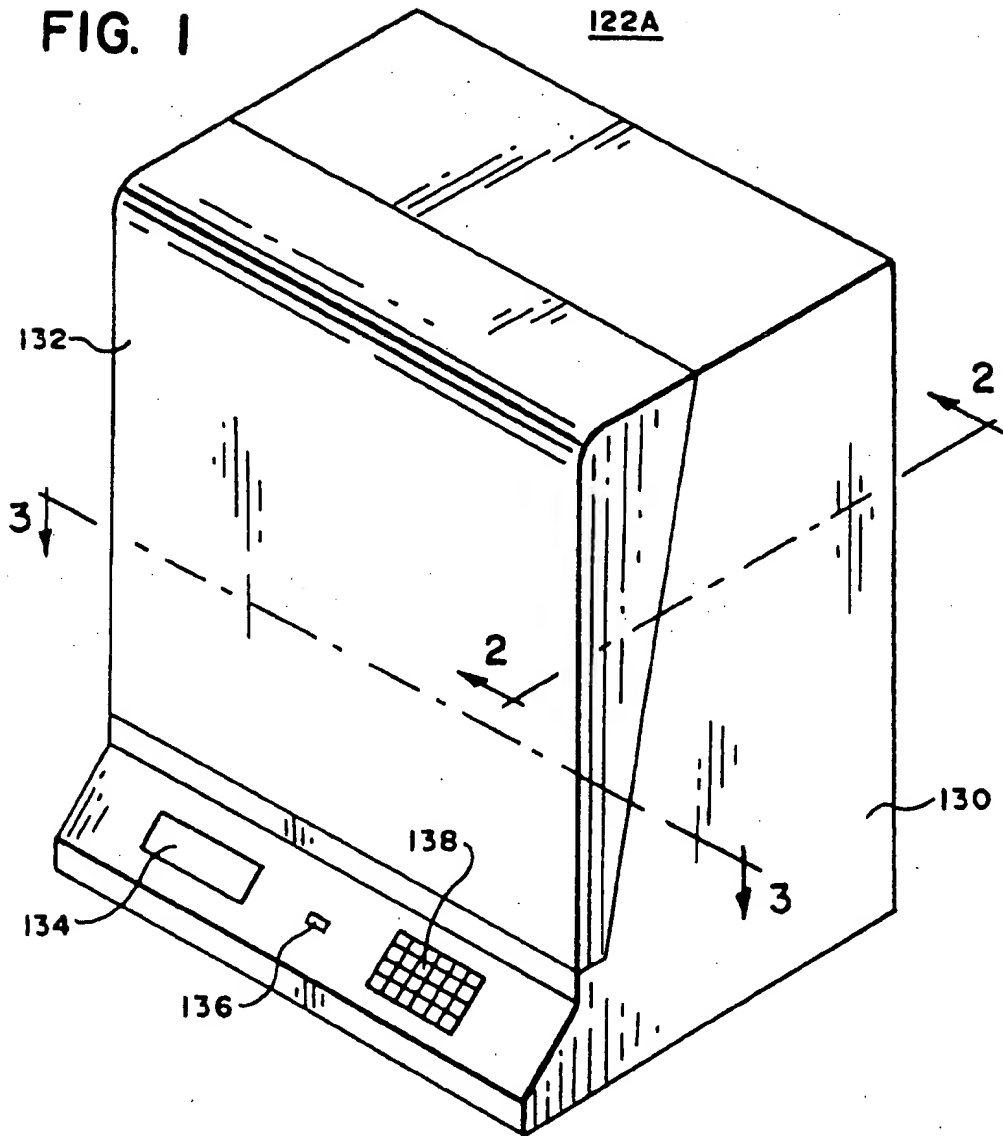
122A

FIG. 2

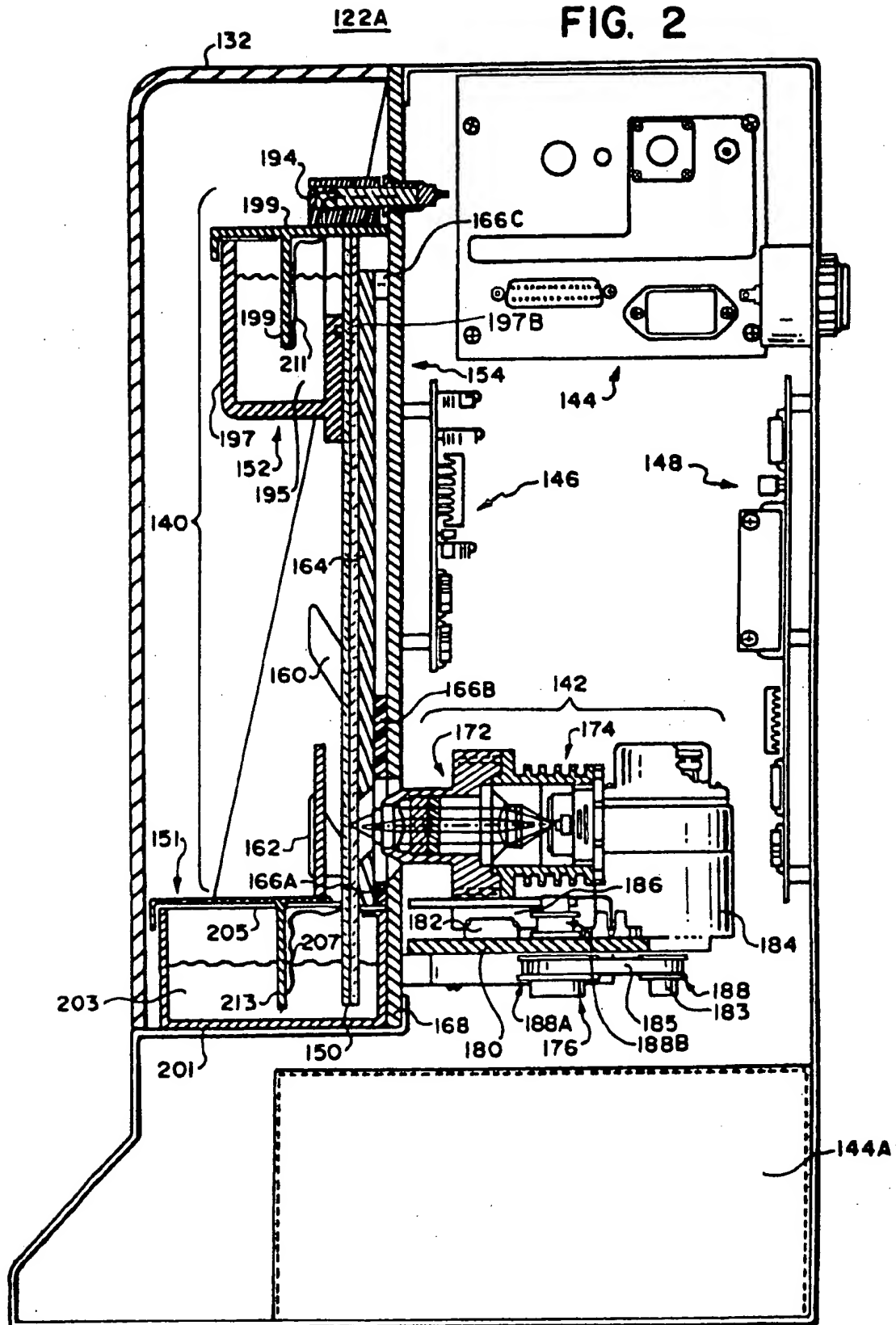


FIG. 3

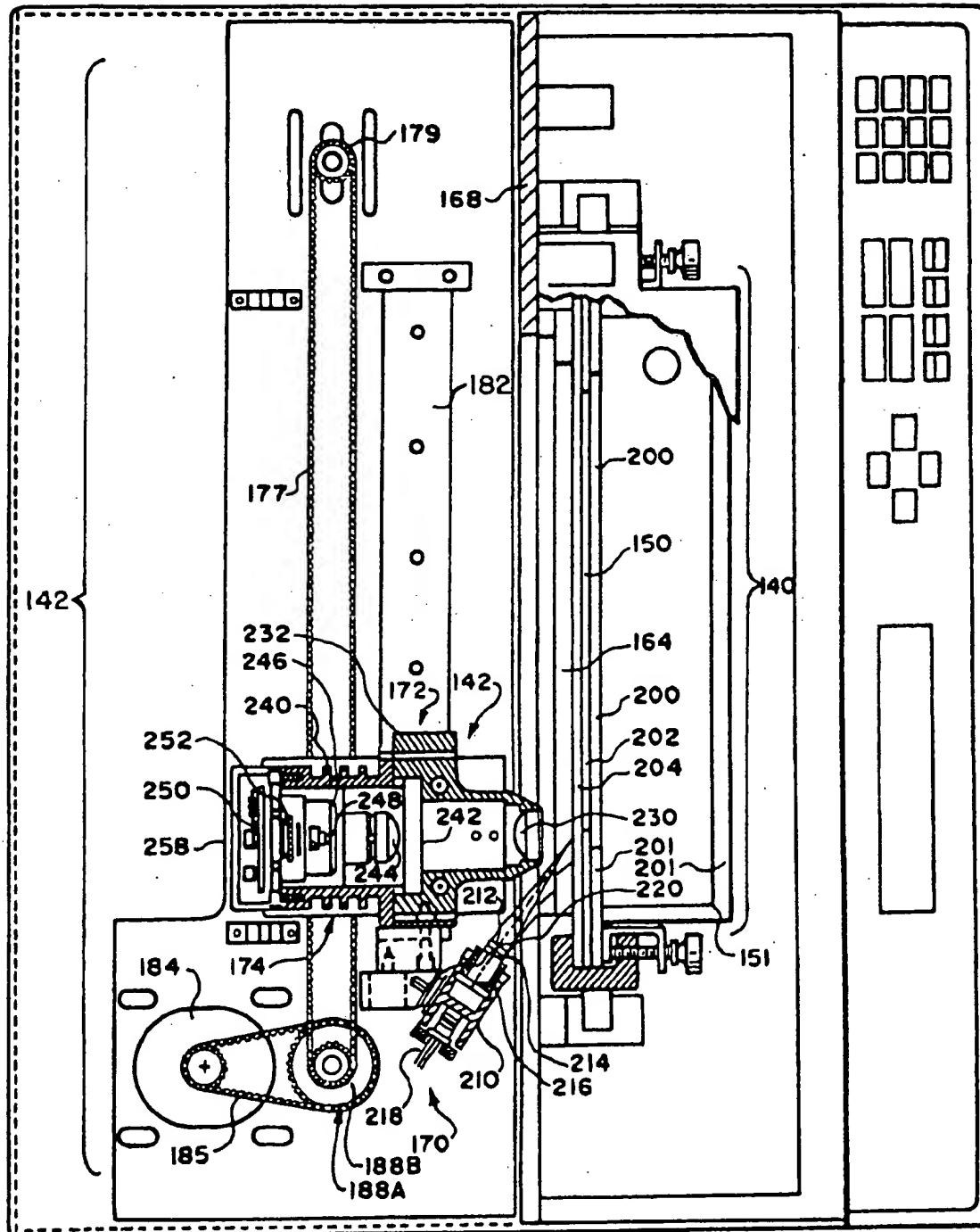


FIG. 4

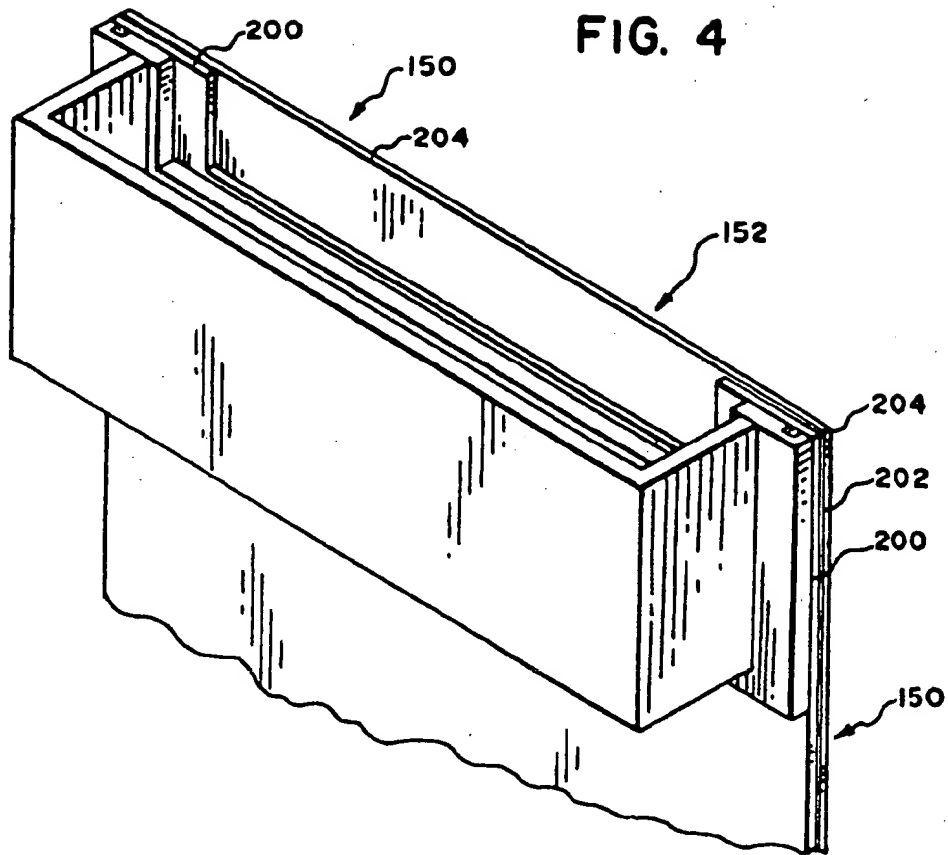


FIG. 5

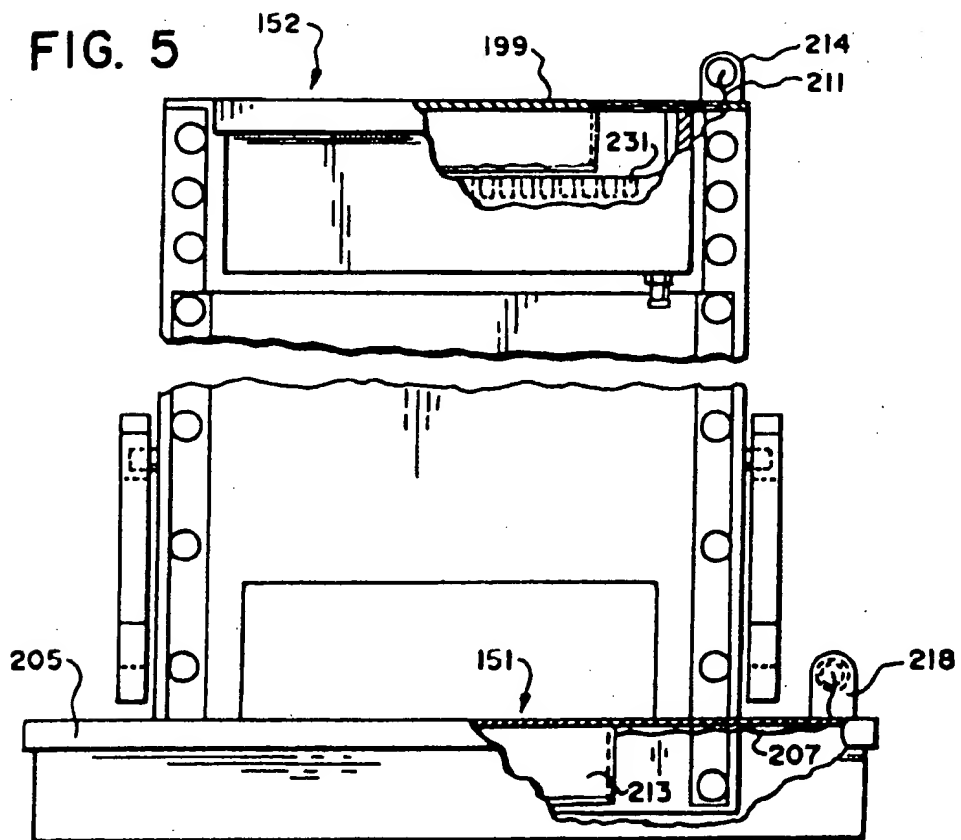
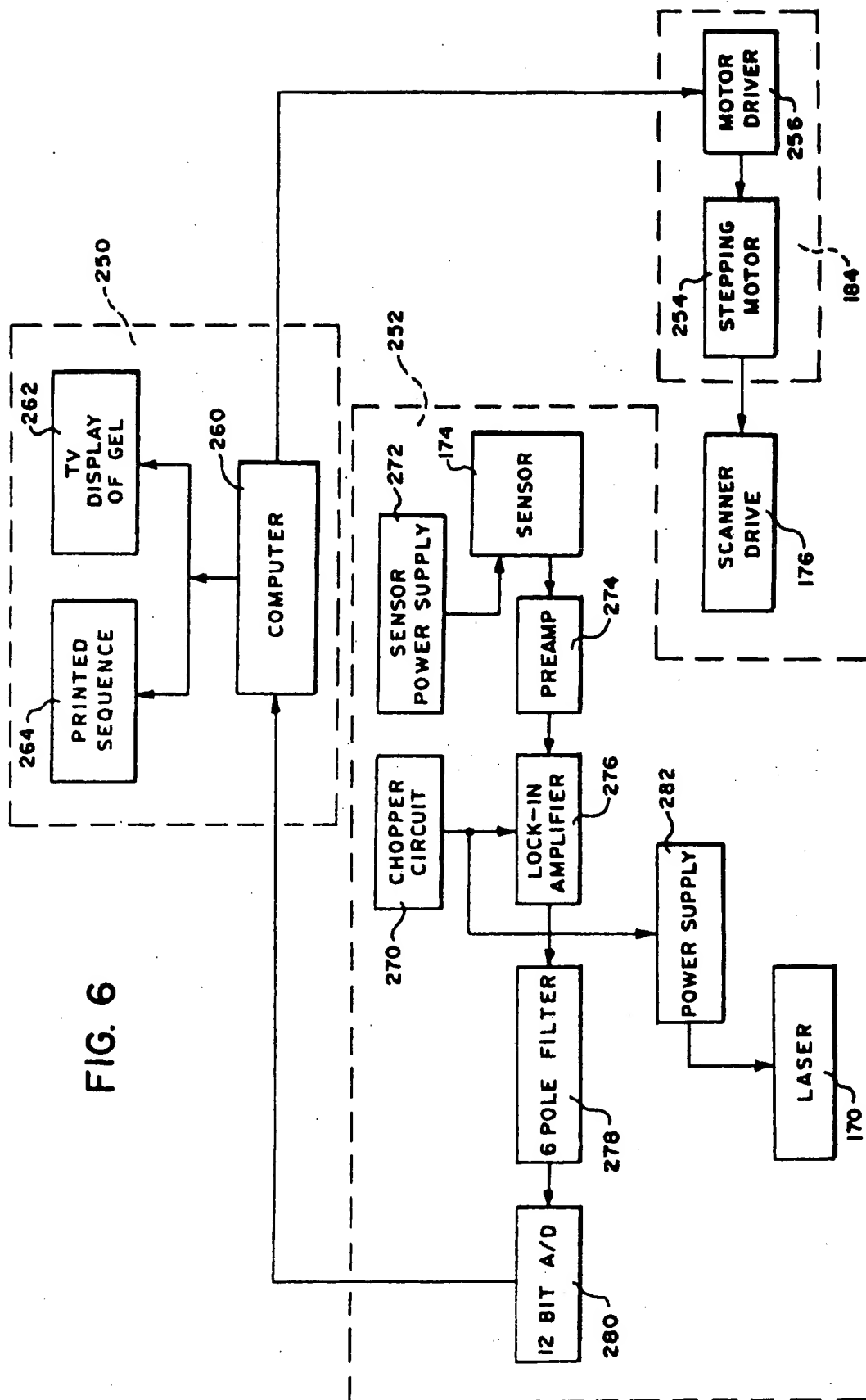


FIG. 6



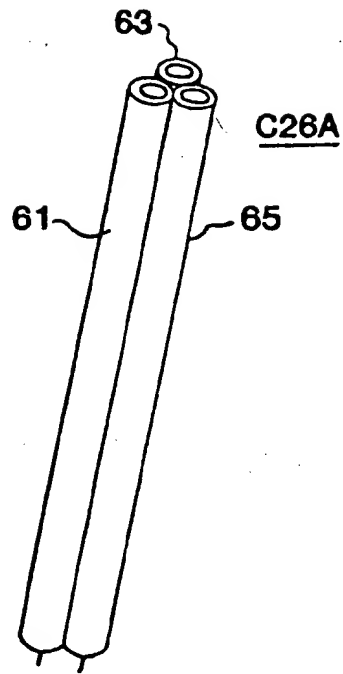


FIG. 7



European Patent  
Office

## EUROPEAN SEARCH REPORT

Application Number

DOCUMENTS CONSIDERED TO BE RELEVANT			EP 96250149.0
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int. Cl. 6)
E	EP - A - 0 670 374 (LI-COR, INC.) * Claims *	1-4	C 12 Q 1/68 C 07 D 209/62 C 09 B 23/00
A	EP - A - 0 533 302 (LI-COR, INC.) * Claims 1-5 *	1-3	
A	WO - A - 95/07 361 (INSTITUT PASTEUR et al.) * Claims *	1	
			TECHNICAL FIELDS SEARCHED (Int. Cl. 6)
			C 12 Q C 07 D C 09 B
The present search report has been drawn up for all claims			
Place of search VIENNA		Date of completion of the search 09-10-1996	Examiner WOLF
<p><b>CATEGORY OF CITED DOCUMENTS</b></p> <p>X : particularly relevant if taken alone  Y : particularly relevant if combined with another document of the same category  A : technological background  O : non-written disclosure  P : intermediate document</p> <p>T : theory or principle underlying the invention  E : earlier patent document, but published on, or after the filing date  D : document cited in the application  L : document cited for other reasons</p> <p>&amp; : member of the same patent family, corresponding document</p>			

EPO FORM 1503 03.82 (P0401)